

Les systèmes d'information au service de la surveillance épidémiologique : enjeux et défis liés à la pandémie de Covid-19 // Information systems for epidemiological surveillance: Issues and challenges related to the Covid-19 pandemic

Coordination scientifique // Scientific coordination

Perrine de Crouy-Chanel et **Julien Durand** (Santé publique France, Saint-Maurice)

Et pour le Comité de rédaction du BEH : **Franck De Laval** (Cespa, Marseille), **Matthieu Eveillard** (Université d'Angers), **Marie-Pierre Tavolacci** (CHU de Rouen), **Bertrand Gagnière** (Santé publique France – Bretagne, Rennes), **Arnaud Tarantola** (Santé publique France – Ile-de-France, Saint-Denis)

> SOMMAIRE // Contents

ÉDITORIAL // Editorial

Les systèmes d'information face à la pandémie de Covid-19 : défis majeurs et héritages
// Information systems in the face of the COVID-19 pandemic: Major challenges and legaciesp. 438

Laëtitia Huiart, Caroline Alleaume

Université de Versailles Saint-Quentin-en-Yvelines – Paris-Saclay, Assistance publique des Hôpitaux de Paris, directrice scientifique de Santé publique France de 2020 à 2024

ARTICLE // Article

Surveillance virologique en France : impact de la Covid-19 et perspectives
// Virological surveillance in France: The impact of COVID-19 and future prospectsp. 440

Julien Durand et coll.

Santé publique France, Saint-Maurice

FOCUS // Focus

La surveillance épidémiologique en temps proche du réel des hospitalisations liées à la Covid-19 en France, 2020-2023 : SI-VIC
// Near real-time epidemiological surveillance of COVID-19-related hospital admissions in France, 2020-2023: The SI-VIC surveillance systemp. 450

Yann Le Strat et coll.

Santé publique France, Saint-Maurice

ARTICLE // Article

VAC-SI : un système d'information pour le suivi de la couverture vaccinale des vaccins contre la Covid-19
// VAC-SI: A surveillance system to monitor COVID-19 vaccination coveragep. 454

Johnny Platon et coll.

Santé publique France, Saint-Maurice

ARTICLE // Article

Surveillance de la pandémie de Covid-19 : contribution et performances du système SurSaUD®
// Surveillance of the COVID-19 pandemic: Contribution and performance of the SurSaUD® systemp. 462

Anne Fouillet et coll.

Santé publique France, Saint-Maurice

ARTICLE // Article

Santé publique France face au défi de l'open data
// Santé publique France rises to the challenge of open datap. 475

Étienne Lucas et coll.

Santé publique France, Saint-Maurice

(Suite page 438)

La reproduction (totale ou partielle) du BEH est soumise à l'accord préalable de Santé publique France. Conformément à l'article L. 122-5 du code de la propriété intellectuelle, les courtes citations ne sont pas soumises à autorisation préalable, sous réserve que soient indiqués clairement le nom de l'auteur et la source, et qu'elles ne portent pas atteinte à l'intégrité et à l'esprit de l'œuvre. Les atteintes au droit d'auteur attaché au BEH sont passibles d'un contentieux devant la juridiction compétente.

Retrouvez ce numéro ainsi que les archives du Bulletin épidémiologique hebdomadaire sur <https://www.santepubliquefrance.fr/revues/beh/bulletin-epidemiologique-hebdomadaire>

Directrice de la publication : Caroline Semaille, directrice générale de Santé publique France
Rédactrice en chef : Valérie Colombani-Cocuron, Santé publique France, redaction@santepubliquefrance.fr
Rédactrice en chef adjointe : Frédérique Biton-Debernardi
Responsable du contenu en anglais : Chloé Chester
Secrétariat de rédaction : Quentin Lacaze
Comité de rédaction : Raphaël Andler, Santé publique France ; Thomas Bénét, Santé publique France - Auvergne-Rhône-Alpes ; Florence Bodeau-Livinec, EHESP ; Kathleen Chamli, Santé publique France ; Perrine de Crouy-Chanel, Santé publique France ; Olivier Dejardin, CHU Caen ; Franck de Laval, Cespa ; Martin Herbas Ekat, CHU Brazzaville, Congo ; Matthieu Eveillard, CHU Angers ; Bertrand Gagnière, Santé publique France - Bretagne ; Isabelle Grémy ; Anne Guinard, Santé publique France - Occitanie ; Camille Lecoffre-Bernard, Santé publique France ; Élodie Lebréton, Santé publique France ; Yasmina Ouharroune, Santé publique France ; Valérie Ollé, Santé publique France ; Arnaud Tarantola, Santé publique France - Ile-de-France ; Marie-Pierre Tavolacci, CHU Rouen ; Hélène Therre, Santé publique France ; Isabelle Villena, CHU Reims ; Marianne Zeller, IFR Sciences de santé de Dijon.
Santé publique France - Site Internet : <https://www.santepubliquefrance.fr>
Préresse : Luminess
ISSN : 1953-8030

ARTICLE // Article

SurvESMS : un dispositif de surveillance de la Covid-19 en établissements sociaux et médico-sociaux
// SurvESMS: A system for monitoring COVID-19 in long-term care facilitiesp. 481

Côme Daniau et coll.
Santé publique France, Saint-Maurice

FOCUS // Focus

Les systèmes d'information à l'épreuve de la Covid-19 : enseignements, nouveaux enjeux et perspectives pour se préparer aux prochaines crises
// Information systems put to the test by COVID-19: Lessons learned, new challenges and opportunities to prepare for future crisesp. 488

Céline Caserio-Schönemann et coll.
Santé publique France, Saint-Maurice

ÉDITORIAL // Editorial

LES SYSTÈMES D'INFORMATION FACE À LA PANDÉMIE DE COVID-19 : DÉFIS MAJEURS ET HÉRITAGES

// INFORMATION SYSTEMS IN THE FACE OF THE COVID-19 PANDEMIC: MAJOR CHALLENGES AND LEGACIES

Laëtitia Huiart¹, Caroline Alleaume²

¹ Université de Versailles Saint-Quentin-en-Yvelines – Paris-Saclay, Assistance publique des Hôpitaux de Paris, directrice scientifique de Santé publique France d'octobre 2020 à août 2024

² Santé publique France, Saint-Maurice

L'émergence de la pandémie de Covid-19 en 2020 a mis en lumière, de manière inédite, la nécessité de disposer de systèmes d'information robustes, réactifs et adaptés pour surveiller une possible émergence et la propagation du virus, ses caractéristiques et conséquences, et fournir des indicateurs de santé publique en temps quasi-réel. L'importance de ces systèmes réside, non seulement dans leur capacité à suivre l'évolution de la pandémie, mais aussi à contribuer à l'évaluation des mesures de contrôle mises en place et ainsi à guider les réponses politiques, sanitaires et sociales dans un contexte de crise mondiale. Dans l'urgence de la pandémie, ces systèmes d'information ont initialement dû mobiliser les données disponibles, puis s'adapter avec agilité, afin de fournir de nouveaux indicateurs plus précis et pertinents.

Plusieurs défis majeurs ont dû être relevés : 1) collecter et analyser de vastes quantités de données quasi-exhaustives pour décrire la circulation virale au niveau territorial ; 2) fournir des indicateurs fiables et actualisés quotidiennement, afin de permettre une prise de décision rapide et éclairée ; 3) intégrer une vision longitudinale des données, nécessitant un chaînage entre les données issues des différents systèmes, permettant par exemple de suivre les multiples tests effectués par chaque individu, d'étudier les liens entre l'infection et la gravité de la maladie en fonction des variants impliqués ou du statut vaccinal.

L'un des principaux défis dans la gestion de la pandémie de Covid-19 a été la mise en place de systèmes capables de capter fidèlement l'évolution de l'épidémie dans son ensemble. Les acteurs de l'écosystème de la santé ont ainsi œuvré ensemble à la mise en place, dès 2020, d'un système de surveillance multisource. Les dispositifs créés *ex nihilo*, tels que le système

de surveillance des résultats des tests biologiques, celui de suivi du déploiement de la campagne vaccinale ou encore celui de surveillance de l'activité hospitalière (ce dernier ayant été adapté à la crise), ont permis la production d'indicateurs à la fois précis et déclinables à une échelle territoriale très fine, facilitant ainsi la mise en œuvre de mesures adaptées. La collecte d'informations quasi-exhaustives est une caractéristique commune des systèmes de surveillance développés pendant la crise, ainsi que leur agilité pour s'adapter à l'évolution des politiques publiques de dépistage ou de vaccination. Cinq articles de ce numéro¹⁻⁵ présentent chacun de ces dispositifs et discutent les enjeux de leur déploiement en urgence. Ils montrent à quel point les acteurs impliqués dans cette surveillance ont été confrontés à des défis techniques nouveaux, notamment en raison des volumes considérables de données à collecter et à traiter quotidiennement. Malgré ces défis, cette approche s'est avérée essentielle pour garantir une réponse cohérente et adaptée à l'évolution de la pandémie.

La dimension inédite et la dynamique évolutive de la pandémie Covid-19 ont rendu cruciale la rapidité d'accès à de l'information fiable. Les systèmes de surveillance reposant sur la capacité de déclarants à saisir des données de manière réactive ont rapidement été mis à rude épreuve. Cela a constitué une difficulté majeure dans l'analyse des indicateurs produits et leur restitution possible. Les systèmes d'information automatisés ont dû intégrer cette dimension temporelle dans la collecte et la restitution des indicateurs pour fournir aux décideurs des actualisations quotidiennes. Les plateformes numériques de déclaration des résultats de tests virologiques ont ainsi presque instantanément absorbé les données dans des systèmes

centralisés. L'usage croissant de systèmes de collecte, d'analyse et de restitution automatisés a joué un rôle central dans cette réactivité. Le traitement de données et la restitution des résultats en temps quasi-réel sont en effet un enjeu de transparence et de fiabilisation de la parole publique. Ils permettent à l'ensemble des citoyens d'accéder à une information de référence, utile à la compréhension de la situation pour guider l'adaptation des comportements individuels. Ainsi, Santé publique France, dans le cadre d'une stratégie française nationale volontariste de partage de l'information, a progressivement rendu publics plus de 150 indicateurs par jour, produit des analyses qualitatives hebdomadaires, et a pu accompagner l'évolution des besoins en connaissances tout au long de l'épidémie. Un article de ce numéro décrit cette montée en charge de l'open data⁶ permettant la visualisation des indicateurs, mais aussi leur réutilisation par différents acteurs du monde de la recherche, de la presse ou de plateformes spécifiques de datavisualisation. Ces indicateurs ont notamment alimenté les travaux des équipes de modélisation pour anticiper l'évolution de la circulation virale, ses répercussions et estimer l'effet des mesures de santé publique. Cette information a permis aux décideurs des différents échelons territoriaux (régionaux, nationaux et internationaux), mais aussi aux acteurs, tels que les hôpitaux, d'ajuster leurs réponses de façon réactive : anticiper les pics épidémiques, renforcer les mesures de confinement, adapter les campagnes de vaccination, ou encore réorganiser les services de soins intensifs.

La capacité à interconnecter les données issues des différents systèmes d'information et la prise en compte de la dimension longitudinale de ces données ont été des enjeux importants pour une meilleure compréhension de l'épidémie. Le suivi du devenir et des recours aux soins des individus testés positifs a permis d'estimer la gravité de la maladie, et de la réévaluer à chaque émergence de variant. Les données produites par le système de surveillance de la campagne vaccinale, couplées à celles des tests et des hospitalisations, ont permis d'évaluer rapidement l'effet des vaccins et leur évolution au fil du temps. Le chaînage a été complexifié par l'absence d'identifiant unique de chaque patient, nécessitant le recours à la pseudonymisation. Ce recours complique le travail et peut en altérer la fiabilité. Les dispositifs réglementaires devront à l'avenir être adaptés pour faciliter l'utilisation d'identifiants favorisant l'interconnexion des bases de données. Ceci permettra de répondre plus efficacement aux grands enjeux de santé publique soulevés par une crise sanitaire de cette envergure. Le dernier article de ce numéro⁷ souligne d'ailleurs que les systèmes d'information à l'ère post-Covid doivent être interopérables.

Notre capacité à capitaliser sur les systèmes d'information déployés pendant la crise de la Covid-19 est aussi cruciale pour moderniser plus généralement nos systèmes de surveillance de routine. La pérennisation de ces systèmes, en particulier de ceux permettant l'accès aux données de laboratoire ou aux patients hospitalisés, doit donc être envisagée afin d'assurer un héritage de cet extraordinaire déploiement. C'est dans cette perspective que le dernier article conclut

ce numéro⁷, tirant les leçons d'une période de crise riche en développements. Dans cette optique, le projet « Orchidée », visant à organiser un réseau national de centres hospitaliers universitaires impliqués dans la surveillance épidémiologique, est particulièrement intéressant. Ce projet repose sur l'utilisation secondaire des données issues des entrepôts de données hospitaliers pour produire des indicateurs épidémiologiques, afin de répondre aux besoins de la surveillance de routine tout en préparant des réponses aux situations sanitaires exceptionnelles, y compris les émergences de nouveaux virus.

Les répercussions sur la santé de la population de la crise dépassent largement les enjeux strictement liés au virus et vont bien au-delà des conséquences possibles à long terme d'infections individuelles. Ainsi, nos systèmes de surveillance, en particulier les systèmes syndromiques, ont détecté précocement la dégradation de la santé mentale au sein de la population. Les enquêtes, qu'elles soient générales ou spécifiques, ont non seulement mis en lumière l'impact global sur la santé, mais aussi des changements profonds dans les recours aux soins préventifs et les comportements de santé. Cela laisse présager des conséquences à long terme de la crise sur la santé de la population. Une approche globale de la surveillance s'avère donc essentielle pour appréhender l'ensemble des répercussions engendrées par une telle crise. En outre, l'impact a été socialement très différencié, affectant de manière disproportionnée les plus vulnérables. Les données sociales ou économiques individuelles disponibles au sein des systèmes de surveillance sont toutefois très limitées, et la documentation de ces impacts différenciés s'avère complexe et souvent retardée. Seuls les indicateurs géographiques de défavorisation sociale ont pu être élaborés rapidement, grâce à la quasi-exhaustivité des bases de données, facilitant ainsi le ciblage des actions de prévention à des niveaux précis. L'intégration systématique des données, limitant au maximum l'implication de « déclarants », et la capacité à relier les informations socio-économiques constituent des enjeux majeurs pour les systèmes de surveillance à venir.

Enfin, adopter une approche intégrée de la surveillance, coordonnée nationalement et internationalement, favorisera une vision holistique de la santé dans une perspective « Une seule santé », englobant les dimensions humaine, animale et environnementale, et ce, au-delà des frontières. Concevoir une surveillance sous cet angle soulève de nombreux enjeux techniques et réglementaires, mais permettra de répondre aux grands défis de santé publique à venir. ■

Liens d'intérêt

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

[1] Durand J, Fayad M, Feri A, Truong E, Taisne B, Amy E, *et al.* Surveillance virologique en France : impact de la Covid-19 et perspectives. Bull Épidémiol Hebd. 2024;(20-21):440-9 http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_1.html

[2] Le Strat Y, Chereau F. Focus. La surveillance épidémiologique en temps proche du réel des hospitalisations liées au Covid-19 en France, 2020-2023 : SI-VIC. Bull Épidémiol Hebd. 2024;(20-21):450-4. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_2.html

[3] Platon J, Fonteneau L, Hanguelhard R, Gagnière B, Gault G, Deschamps G, *et al.* VAC-SI : un système d'information pour le suivi de la couverture vaccinale des vaccins contre la Covid-19. Bull Épidémiol Hebd. 2024;(20-21):454-61. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_3.html

[4] Fouillet A, Pontais I, Forgeot C, Naud J, Pedrono G, Jezewski-Serra D, *et al.* Surveillance de la pandémie de Covid-19 : contribution et performances du système SurSaUD®. Bull Épidémiol Hebd. 2024;(20-21):462-74. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_4.html

[5] Daniau C, Naud J, Tessier S, Hassan E. SurVESMS : un dispositif de surveillance de la Covid-19 en établissements

sociaux et médico-sociaux. Bull Épidémiol Hebd. 2024;(20-21):481-8. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_6.html

[6] Lucas É, Fouquet A, Jezewski-Serra D, Ben Raies J, de Crouy-Chanel P, Alleaume C. Santé publique France face au défi de l'open data. Bull Épidémiol Hebd. 2024;(20-21):475-81. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_5.html

[7] Caserio-Schönemann C, Arfaoui A, Coignard B, Le Strat Y, Rolland P, Spaccaverri G. Les systèmes d'information à l'épreuve de la Covid-19 : enseignements, nouveaux enjeux et perspectives pour se préparer aux prochaines crises. Bull Épidémiol Hebd. 2024;(20-21):488-90. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_7.html

Citer cet article

Huiart L, Alleaume C. Éditorial. Les systèmes d'information face à la pandémie de Covid-19 : défis majeurs et héritages. Bull Épidémiol Hebd. 2024;(20-21):438-40. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_0.html

ARTICLE // Article

SURVEILLANCE VIROLOGIQUE EN FRANCE : IMPACT DE LA COVID-19 ET PERSPECTIVES

// VIROLOGICAL SURVEILLANCE IN FRANCE: THE IMPACT OF COVID-19 AND FUTURE PROSPECTS

Julien Durand¹ (julien.durand@santepubliquefrance.fr), **Myriam Fayad¹**, **Adeline Feri¹**, **Émilie Truong¹**, **Benjamin Taisne¹**, **Étienne Amy²**, **Philippe Mati²**, **Laura Reibel²**, **Yves Pascal³**, **Achille Lerpinière³**, **Raphaël Beaufret³**, **Bruno Gauthier⁴**, **Catherine Coignard⁵**, **Anne Ebel⁵**, **Bénédicte Roquebert⁶**, **Stéphanie Haïm-Boukoba⁶**, **Éric Perez⁶**, **Shérif Baldé⁶**, **Pascale Bernillon¹**, **Camille Pelat¹**, **Edouard Chatignoux¹**, **Christophe Bonaldi¹**, **Guillaume Spaccaverri¹**, **Adel Arfaoui¹**, **Didier Che¹**, **Céline Caserio-Schönemann¹**, **Bruno Coignard¹**, **Yann Le Strat¹**

¹ Santé publique France, Saint-Maurice

² Assistance publique-hôpitaux de Paris (AP-HP), Paris

³ Ministère de la Santé et de la Prévention, Paris

⁴ Laboratoire Inovie, Montpellier

⁵ Laboratoire Eurofins Biomnis, Lyon

⁶ Laboratoire Cerba, Issy-les-Moulineaux

Soumis le 17.10.2023 // Date of submission: 10.17.2023

Résumé // Abstract

La surveillance épidémiologique est primordiale pour orienter les décideurs vers les mesures de gestion adaptées aux différents contextes d'une crise sanitaire. Dans le cadre de la pandémie liée au SARS-CoV-2, une surveillance épidémiologique a été mise en place, permettant notamment d'orienter la stratégie « tester, alerter, protéger » dans le but de freiner la propagation du virus.

Un système d'information de dépistage populationnel (SI-DEP) a été mis en place le 13 mai 2020 de manière très réactive, deux mois après la disponibilité des tests en mars 2020, pour recevoir systématiquement et d'une façon automatisée tous les résultats de tests (PCR, antigéniques) effectués quotidiennement par l'ensemble des laboratoires publics et privés, ainsi que les professionnels de santé habilités (France hexagonale et départements d'outre-mer).

Cet article décrit l'évolution de la surveillance virologique et la mise en place de SI-DEP pendant la crise sanitaire, depuis les phases de réflexions collaboratives entre les différents acteurs impliqués, le développement du système d'information et ses évolutions, jusqu'à l'exploitation et la diffusion des données et indicateurs produits notamment par Santé publique France.

Tout au long de la période de crise, plusieurs défis ont été rencontrés sur les plans scientifiques, techniques, réglementaires, ainsi que de forts enjeux de communication et de partenariats avec l'ensemble des parties prenantes. De fortes capacités d'adaptations, de cohésion et de résilience ont été nécessaires pour assurer l'ensemble des missions tout au long de cette période.

SI-DEP a montré l'importance d'avoir accès aux données des laboratoires de manière réactive pour assurer une surveillance épidémiologique efficace, et déclinée aux niveaux territoriaux les plus fins.

Epidemiological surveillance provides decision-makers with essential information for implementing appropriate measures in response to a health crisis. In the context of the SARS-CoV-2 epidemic, epidemiological surveillance was set up as part of the strategy to trace, test and protect individuals to help control the spread of the virus.

When testing for SARS-CoV-2 became available in March 2020, the French information system for population-based screening (SI-DEP) was developed within 2 months and became operational on 13 May 2020. SI-DEP was designed to systematically and automatically receive the results of all tests (PCR, antigenic, etc.) performed daily by all public and private laboratories in mainland and overseas France.

This article describes the progression of virological surveillance and the implementation of SI-DEP during the health crisis, from the collaborative reflection phases involving various stakeholders to the development of the information system and its subsequent updates. We also explain how the data collected was exploited and communicated in the form of indicators produced by Santé publique France, the French national health agency.

The health crisis presented a number of scientific, technical and regulatory challenges in relation to surveillance. There were also high expectations in terms of communication and partnerships between stakeholders. A significant degree of adaptability, team cohesion and resilience was required to carry out the diverse tasks that this period demanded.

SI-DEP demonstrated the importance of real-time laboratory data to provide an exhaustive and continuous surveillance in conformity with epidemiological monitoring needs.

Mots-clés : SARS-CoV-2, Surveillance, Laboratoires de biologie médicale, Système d'information

// Keywords: SARS-CoV-2, Surveillance, Medical laboratories, Information system

Introduction

L'apport des données de biologie médicale est essentiel à la surveillance des maladies infectieuses (MI) en France. En ce qui concerne les virus respiratoires, la surveillance virologique s'appuie sur le Centre national de référence des virus des infections respiratoires (CNR-VIR). L'objectif général de ce type de surveillance est de (i) détecter et isoler précocement les virus en circulation, (ii) d'en déterminer les caractéristiques antigéniques, et (iii) d'étudier leurs sensibilités aux antiviraux. Lors de l'émergence de la Covid-19, si le diagnostic était initialement basé sur des critères cliniques et radiologiques (scanner), une technique de diagnostic biologique fiable et sensible a été mise au point dès le 16 janvier 2020 par le CNR-VIR, afin de permettre la confirmation des cas, dont les trois premiers en France sont survenus le 24 janvier 2020¹.

À partir de début février 2020, et l'apparition des premières chaînes de transmission locales, il est devenu nécessaire d'augmenter le volume des tests diagnostiques et la capacité à pouvoir les traiter rapidement, afin de pouvoir faire face à l'augmentation de cas suspects. Dans un premier temps réalisés au CNR-VIR, puis dans un nombre croissant de laboratoires hospitaliers, le nombre de tests a rapidement augmenté avec l'arrivée sur le marché de kits de RT-PCR (*polymerase chain reaction*) ciblant le SARS-CoV-2, permettant un usage élargi à tous les laboratoires de biologie publics et privés (arrêté du 7 mars 2020)². C'est alors le début de la surveillance virologique à partir des données de biologie, qui va s'articuler autour du réseau 3-Labos, complété par la remontée des laboratoires hospitaliers en région.

Le réseau 3-Labos, mis en place par Santé publique France et qui fonctionne depuis 2012, est utilisé

pour la surveillance de 21 maladies infectieuses⁽¹⁾, en particulier les arboviroses. Le réseau 3-Labos permet la remontée automatisée et quotidienne de données individuelles de trois laboratoires spécialisés privés (Eurofins-Biomnis, Cerba, et, depuis 2019, Inovie), qui centralisent les prélèvements que d'autres laboratoires leur sous-traitent et représentent une part importante de l'activité de biologie réalisée hors hôpitaux sur le territoire. Les objectifs de 3-Labos sont le suivi des tendances spatio-temporelles des MI ciblées et la contribution à l'investigation d'épidémies. Utilisé seul ou en complément d'autres dispositifs de surveillance, il est évolutif, donc particulièrement adapté en cas de maladie émergente pour obtenir des données de manière réactive.

Ainsi, la remontée des résultats vers Santé publique France via le système 3-Labos a permis de produire, dès le 14 mars 2020, l'un des premiers indicateurs de circulation du virus, le taux de positivité des tests RT-PCR pour SARS-CoV-2, et d'assurer son suivi dans le temps de façon quotidienne. Bien que ne couvrant qu'une partie des prélèvements réalisés sur le territoire, les données permettaient notamment de décrire les personnes testées et les personnes positives en termes d'âge, de sexe et de répartition géographique. En revanche, il ne permettait pas un suivi de la circulation virale à un niveau territorial fin et n'était pas adapté pour les activités de *contact-tracing* (recherche des contacts associés aux cas), compte tenu de sa couverture limitée et non exhaustive.

⁽¹⁾ Borréliose de Lyme, chikungunya, coqueluche, dengue, fièvre Q, hépatites A et E, papillomavirus humain, infections à gonocoque, chlamydia et *Mycoplasma genitalium*, leptospirose, psittacose, rougeole, rubéole, syphilis, encéphalites à tique, toxoplasmose, tuberculose, West Nile, et Zika.

Ce constat, ainsi que la nécessité pour les pouvoirs publics de disposer d'un système permettant de mettre en lien les professionnels de santé, les personnes testées et les autorités de santé, a rapidement fait émerger le besoin d'un dispositif national centralisant de manière exhaustive, automatisée, systématique et réactive les données de l'ensemble des laboratoires publics et privés.

Cet article décrit l'évolution de la surveillance virologique opérée par Santé publique France pendant la crise sanitaire causée par le SARS-CoV-2, depuis les phases de réflexions collaboratives en février 2020 entre les différents acteurs impliqués, le montage du système d'information et ses évolutions, l'exploitation des données recueillies pour la production d'indicateurs de surveillance, puis la transition vers d'autres systèmes d'information à partir de juillet 2023. Seront également abordés les défis liés aux évolutions successives de ce système pour s'adapter à la pandémie, ainsi que les enjeux liés à la diffusion de données de manière réactive, notamment via l'open data.

Retour d'expérience du système d'information de dépistage populationnel (SI-DEP)

La mise en place d'un nouveau système d'information : SI-DEP

Le régime de l'état d'urgence sanitaire créé par la loi n° 2020-290 du 23 mars 2020 d'urgence pour faire face à l'épidémie de SARS-CoV-2³ a permis de prendre les mesures rendues nécessaires par ces circonstances, dont la création de systèmes d'information nationaux pour la remontée de données de santé. Le Système d'information de dépistage populationnel (SI-DEP) est un système d'information (SI) national sécurisé pour le recueil des résultats de tests de dépistage de SARS-CoV-2 effectués en France.

Mis en production le 13 mai 2020, SI-DEP a été construit dans l'objectif de répondre aux besoins de surveillance épidémiologique à partir des données de biologie pendant la crise sanitaire, et afin d'orienter les mesures de gestion permettant de limiter l'impact de l'épidémie, d'alimenter les outils de *contact-tracing*, de permettre l'accès à des données anonymisées pour la recherche (décret n° 2020-551 du 12 mai 2020⁴), et de constituer un outil d'échange entre les professionnels de santé et les citoyens.

Ce SI est né d'un partenariat entre le ministère de la Santé et de la Prévention, via la Délégation du numérique en santé et la Direction générale de la santé (responsable de traitement et maître d'ouvrage), l'Assistance publique-Hôpitaux de Paris (AP-HP, maître d'œuvre et chef de file d'un consortium), les laboratoires de biologie médicale publics et privés ainsi que leurs prestataires en SI, l'Agence du numérique en santé (ANS), Santé publique France, la Direction de la recherche, des études, de l'évaluation

et des statistiques (Drees), la Caisse nationale d'assurance maladie (Cnam), et les agences régionales de santé (ARS).

SI-DEP a permis le raccordement de près de 5 000 laboratoires privés et publics disposant ou non d'automates pour la réalisation simultanée d'un grand nombre de tests (dits « MGI »), dont 600 plateaux techniques, à un système d'information centralisé pour le recueil des tests RT-PCR. Les laboratoires poursuivent la saisie des données concernant les résultats de tests et les données sociodémographiques dans leur système de gestion de laboratoire (SGL), qui sont automatiquement transmises vers SI-DEP. Les données saisies sont collectées dans une base de donnée nationale qui alimente les différents acteurs pour les missions dont ils ont la charge. Les données nominatives sont transmises à la Cnam et aux ARS pour les missions de *contact-tracing*. Seules les données pseudonymisées sont transmises à Santé publique France et à la Drees pour le suivi épidémiologique, et au Système national des données de santé (SNDS) pour la recherche (figure 1). SI-DEP a également facilité l'échange des résultats de tests entre les professionnels de santé et les patients, ainsi que la génération de certificats de tests avec des QR codes pour les passes sanitaires pendant les périodes de confinement et mesures de restrictions nationales.

Un système d'information en constante évolution

Afin de permettre une surveillance capable de s'adapter en permanence à l'épidémie, plusieurs évolutions majeures de SI-DEP ont été nécessaires. En effet, lors de sa mise en place en mai 2020, SI-DEP permettait uniquement la remontée des résultats de RT-PCR pour le dépistage du SARS-CoV-2. Cependant, au cours de la pandémie, la stratégie de tests a évolué de sorte à être toujours plus réactive pour mettre en place de nouvelles mesures de gestion de l'épidémie.

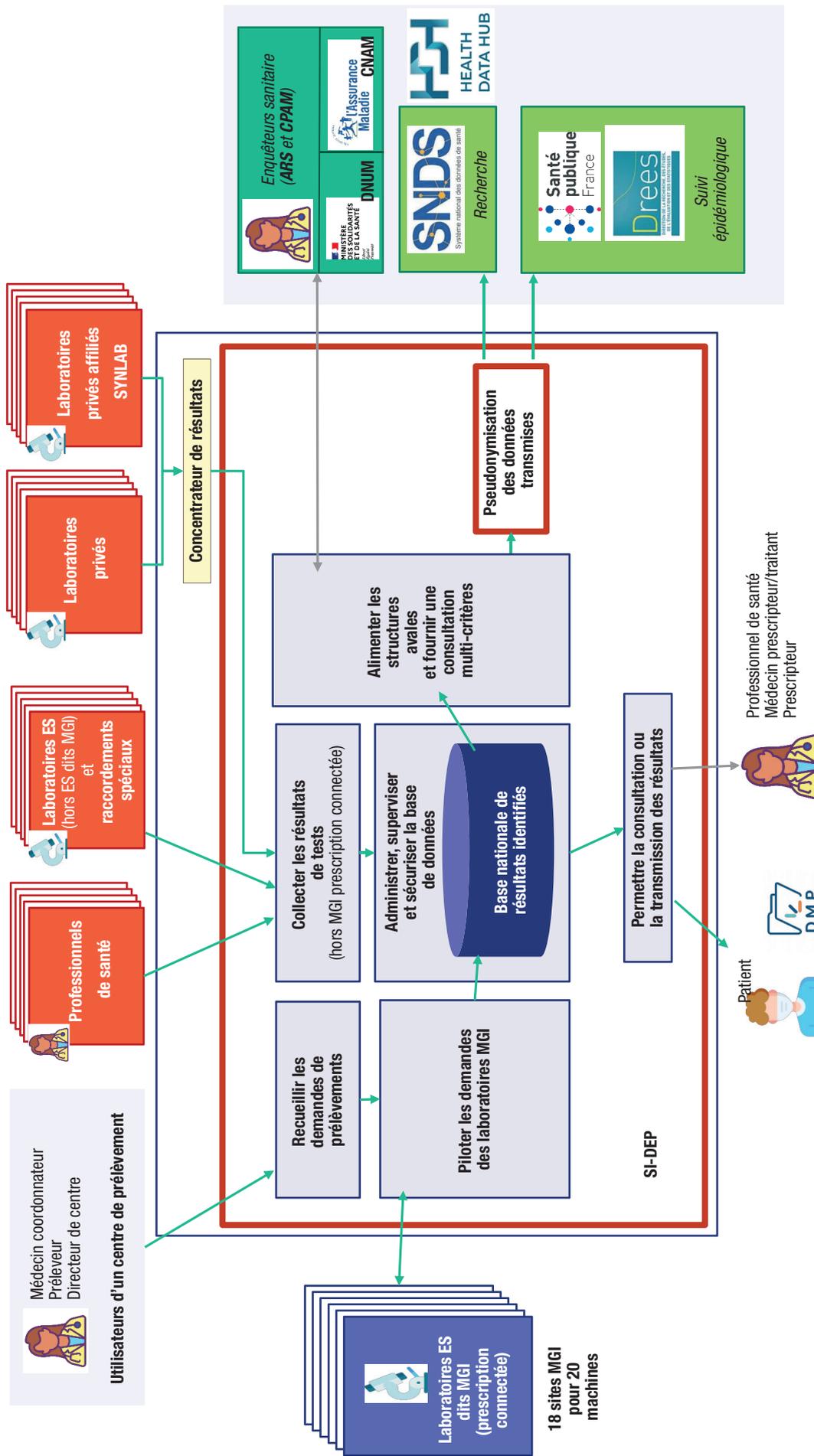
Les changements majeurs ont concerné l'intégration successive : des résultats de tests sérologiques en juillet 2020 ; des résultats de tests antigéniques (TAG) effectués par différents professionnels de santé (pharmaciens, médecins, masseurs-kinésithérapeutes, infirmiers, chirurgiens-dentistes ou encore sages-femmes) en novembre 2020 ; des résultats de criblage et de séquençage pour le suivi des variants en février 2021 puis de certaines mutations ; ainsi que des résultats des autotests supervisés (figure 2).

L'intégration des résultats des TAG dans SI-DEP a notamment pu être effectuée à travers la mise en place d'un portail de saisie connecté à SI-DEP pour les professionnels de santé concernés. Pour faciliter l'exploitation des données saisies manuellement, des contrôles qualité ont été mis en place, afin de garantir une qualité des données comparable à celles remontées par les laboratoires.

Le volume des tests RT-PCR et TAG saisis quotidiennement et la stratégie de dépistage ont aussi évolué au cours du temps, nécessitant une adaptation

Figure 1

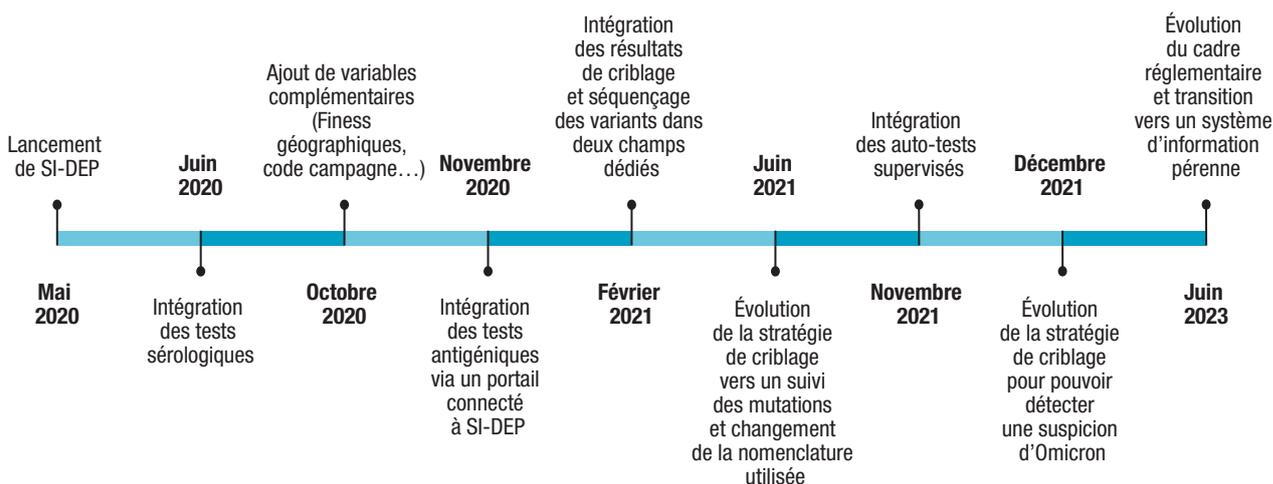
Schéma simplifié de la transmission des données vers et depuis SI-DEP



SI-DEP : Système d'information de dépistage populationnel ; ES : établissements de santé ; MGI : établissements de santé ; CPAM : Caisse primaire d'assurance maladie ; SNDS : Système national des données de santé ; Drees : Direction de la recherche, des études, de l'évaluation et des statistiques.
 Source : AP-HP, équipe SI-DEP.

Figure 2

Les évolutions de SI-DEP entre mai 2020 et juillet 2023



SI-DEP : Système d'information de dépistage populationnel ; Finess : fichier national des établissements sanitaires et sociaux.

continue de l'infrastructure informatique de SI-DEP et une augmentation de ses capacités machine afin de pouvoir absorber la volumétrie croissante de tests réalisés, sans engendrer de problème de transmission de données entre les différentes parties concernées. Lors du lancement de SI-DEP, 40 000 tests par jour étaient enregistrés dans la base. Ce volume a augmenté à 500 000 en mars 2021, puis 1 million en août 2021 pour atteindre jusqu'à 2,7 millions de tests par jour en janvier 2022⁽²⁾, nécessitant une adéquation du dimensionnement des infrastructures techniques tout au long de la crise (la capacité maximale de SI-DEP ayant été portée à 4 millions de tests par jour).

La construction des indicateurs de surveillance virologique

Les flux de données en temps réel issus de l'ensemble des laboratoires et des autres professionnels de santé vers SI-DEP ont rendu possible la production quotidienne d'un ensemble d'indicateurs épidémiologiques déclinés à un niveau territorial très fin, qui ont permis de décrire l'évolution spatio-temporelle de l'épidémie. Le taux d'incidence (nombre de personnes nouvellement testées positives rapporté à la taille de la population), le taux de positivité (nombre de personnes nouvellement testées positives rapporté au nombre total de personnes testées) et le taux de dépistage (nombre de personnes testées rapporté à la taille de la population) sont les indicateurs clés calculés à partir des données issues de SI-DEP⁽³⁾ (figure 3 et 4). Ces indicateurs sont notamment basés sur les résultats des différents tests biologiques (RT-PCR, TAG) et

les caractéristiques individuelles des personnes (âge, sexe, lieu de résidence, présence de symptômes, etc.).

Le calcul des patients incidents – au regard de la définition de cas retenue, selon les recommandations de la Haute Autorité de santé (HAS), du Haut Conseil de la santé publique (HCSP), du Centre européen de prévention et de contrôle des maladies (ECDC) et de l'Organisation mondiale de la santé (OMS), qui décrit un délai de 60 jours par rapport à une précédente infection – a été rendu possible grâce au principe de pseudonymisation. Les informations nominatives d'une personne (nom, prénom, date de naissance et sexe) sont utilisées dans un algorithme qui permet de produire un code unique (pseudonyme) pour chaque personne, donc de repérer des infections successives pour une personne donnée.

Ce pseudonyme contribue à la sécurisation et l'anonymisation des données transmises aux autorités sanitaires, et permettent également d'effectuer des appariements entre les différentes bases de données en lien avec le SARS-CoV-2, comme par exemple les bases de données de vaccination (VAC-SI) et d'hospitalisation (SI-VIC).

Lors de la collecte des informations individuelles du patient au moment de la réalisation d'un test Covid-19, le laboratoire recueille l'adresse de la personne (numéro, voie, code postal et nom de la localité de résidence). Cette adresse complète fait l'objet d'un processus de géocodage (affectation de coordonnées géographiques longitude et latitude) utilisé pour calculer l'Iris (îlot regroupé pour l'information statistique – brique de base en matière de diffusion de données infra-communales⁽⁵⁾) de résidence de la personne testée. Ce processus a permis de décliner des indicateurs à des échelons géographiques très fins (Iris, commune, établissements publics de coopération intercommunale – EPCI), et contribue à l'anonymisation des données envoyées aux autorités sanitaires. La précision de ces indicateurs

⁽²⁾ Santé publique France. Géodes – Nombre de tests réalisés – Quotidien. 2024. https://geodes.santepubliquefrance.fr/#c=indicator&i=sp_capa_quot.t&s=2022-05-01&t=a01&view=map2

⁽³⁾ Santé publique France. Géodes – Covid-19 – Données de laboratoires (SI-DEP – à partir du 23/08/2023). 2024. https://geodes.santepubliquefrance.fr/#c=indicator&i=sp_7j_nt.tx_ti_gliss&s=2023-08-23-2023-08-29&t=a01&view=map2

Figure 3

Taux d'incidence et taux de dépistage, de la semaine du 13 au 19 mai 2020 à la semaine du 25 juin 2023 au 01 juillet 2023, France entière (données au 30 juin 2023)

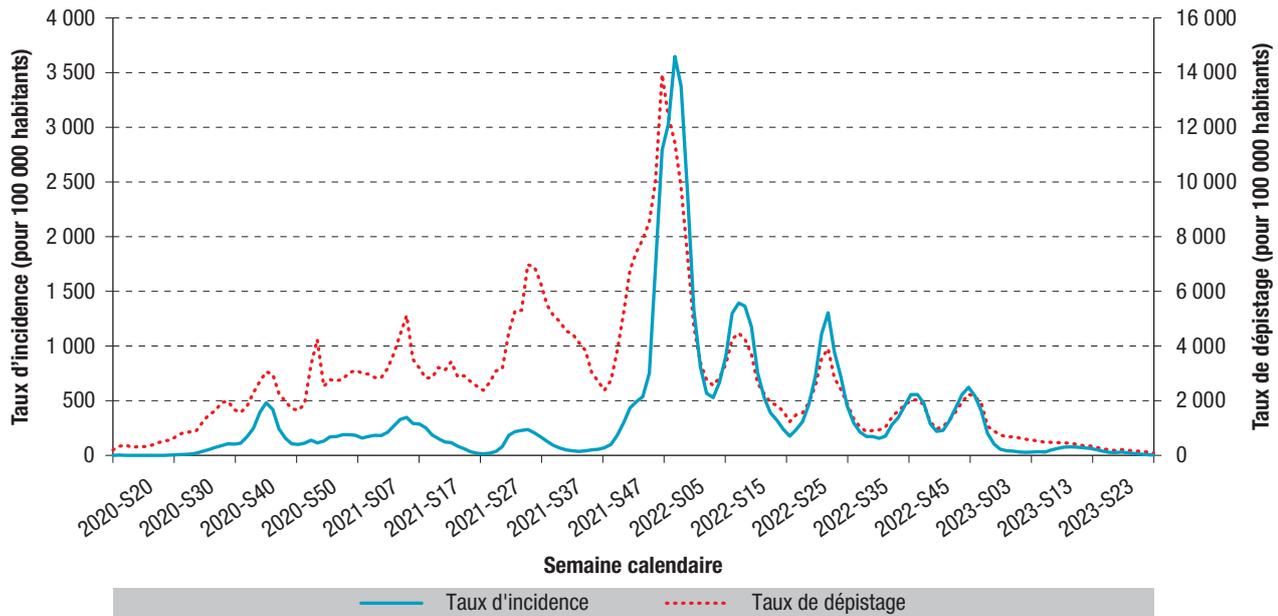
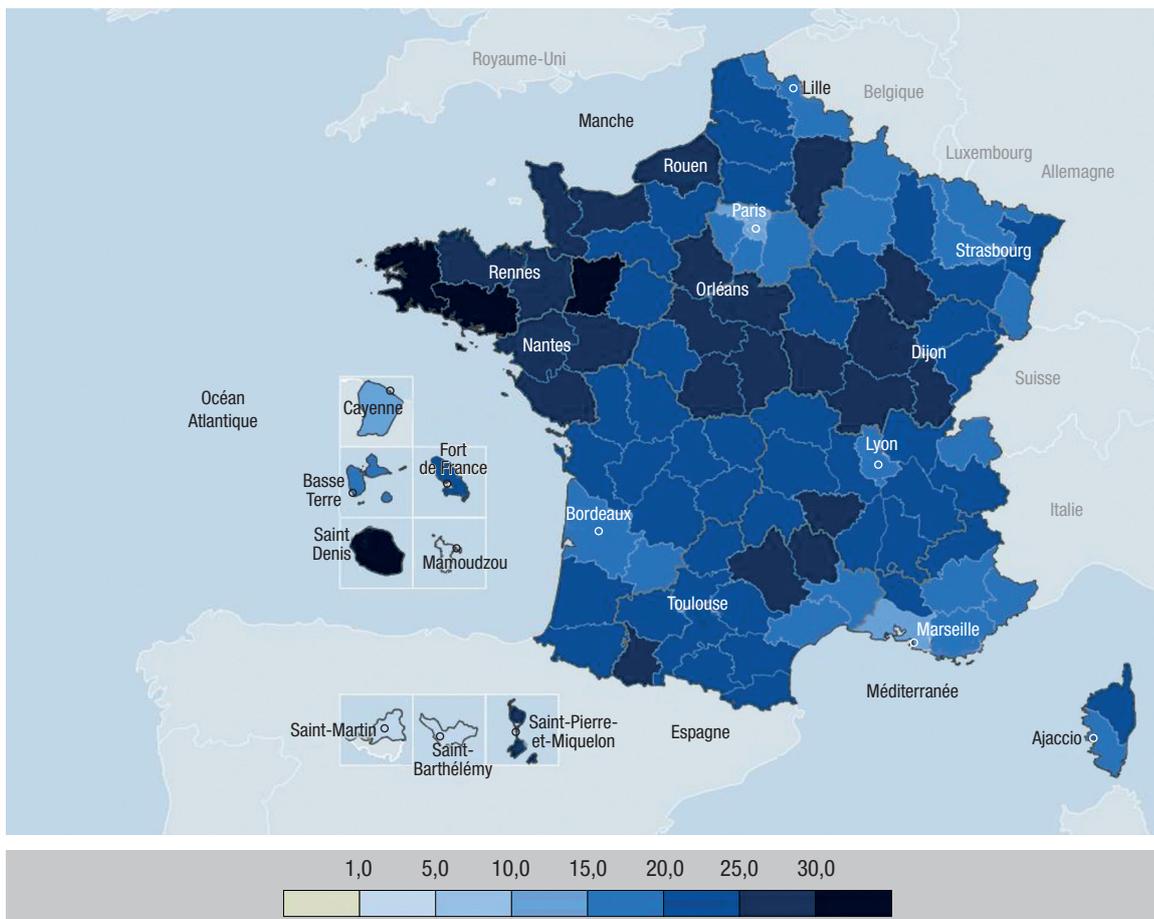


Figure 4

Taux de positivité à la Covid-19 par département pour la semaine 18-2022, France (données au 11 mai 2022)



Taux de positivité - Hebdomadaire - tous âges, 2022-S18 (%) - Source : Base de données SI-LAB issue de SI-DEP.

a permis de décrire la dynamique de l'épidémie et de constituer une aide à la décision par les autorités locales, afin d'adapter les mesures de gestion au plus près des territoires.

L'utilisation des indicateurs issus de SI-DEP pour la gestion de l'épidémie

Les indicateurs issus de SI-DEP en complément des autres sources de données ont contribué à la prise de décision et à la mise en place de mesures de gestion pour faire face à la propagation du virus. Les seuils d'alerte sur le taux d'incidence par exemple, définis pour le suivi de l'évolution de la circulation virale, ont également été utiles pour la déclinaison des différentes mesures sanitaires par les pouvoirs publics (confinement, port du masque, couvre-feu, vaccination...). Ils ont contribué à l'identification des différents pics de l'épidémie, ainsi que des périodes de recrudescence des cas, et à l'adaptation des mesures de freinage, en fonction de la situation sanitaire dans un territoire donné⁶. Par exemple, en octobre 2020, lors de la deuxième vague de Covid-19 et à la suite de l'analyse des indicateurs issus de l'ensemble des sources de surveillance, un couvre-feu a été instauré dans plusieurs métropoles pour freiner la propagation du virus. À la suite de la dégradation de la situation sanitaire sur tout le territoire, cette mesure a évolué vers un confinement national de fin octobre jusqu'à mi-décembre 2020⁶.

En début d'année 2021, à la suite de l'émergence de variants du SARS-CoV-2, la mise en œuvre de RT-PCR de criblage dans les laboratoires a permis de détecter certaines des mutations caractéristiques et d'assurer le suivi de leur diffusion. En complément, les capacités de séquençage ont été progressivement renforcées via le consortium pour la surveillance et la recherche sur les infections à pathogènes émergents via la génomique microbienne (Emergen), coordonné par Santé publique France et l'ANRS – Maladies infectieuses émergentes (ANRS-MIE)⁽⁴⁾. Afin de suivre au niveau national les détections de certains variants connus et considérés comme ayant un impact important en santé publique (plus grande transmissibilité, échappement immunitaire...), le SI-DEP s'est adapté en incluant de nouvelles données saisies par les laboratoires : un premier champ « joker » pour le suivi des résultats de RT-PCR de criblage et un second pour le suivi des résultats de séquençage. L'exploitation de ces champs a permis de produire de nouveaux indicateurs épidémiologiques (taux de criblage, taux de détection d'une mutation donnée) et une estimation très réactive de la proportion de variants porteurs de ces mutations. Elle a aussi permis, dans les premières semaines d'introduction de ces variants, de mettre en place des actions spécifiques de *contact-tracing* dans l'objectif de contenir leur diffusion.

Alors qu'en début d'année 2021, seuls les variants Alpha, Beta et Gamma étaient classés comme préoccupants⁷, le nombre croissant de variants a nécessité, en fin d'année,

l'évolution de l'exploitation des résultats des RT-PCR de criblage vers un suivi des mutations d'intérêt, pour pouvoir suivre l'évolution spatio-temporelle de la proportion des infections dues à un virus porteur de ces mutations.

Les indicateurs virologiques ont également permis de suivre la propagation du SARS-CoV-2 dans certaines populations spécifiques. À la suite de la mise en place de la stratégie de dépistage réactif en milieu scolaire en automne 2021, les indicateurs issus de SI-DEP ont ainsi été déclinés par classe d'âge représentant les différents cycles scolaires (0-2 ans, 3-5 ans, 6-10 ans, 11-14 ans, 15-17 ans, 18 ans et plus)⁽⁵⁾. Ils ont favorisé la mise en place d'aménagements spécifiques en fonction du taux d'incidence chez les élèves pour éviter la fermeture systématique des classes en présence de cas probable ou confirmé⁸.

Les défis techniques et la sécurité des systèmes d'information

La connexion à SI-DEP de l'ensemble des laboratoires réalisant des tests de dépistage du SARS-CoV-2 dans des délais très contraints, entre les mois de mars et mai 2020, a été source de nombreux défis techniques.

Afin de mettre en place une base de données centralisée qui permette la réception de l'ensemble des résultats de tests issus des laboratoires privés (plus de 4 100 laboratoires) et publics (plus de 400 laboratoires), un important travail d'homogénéisation et d'interopérabilité entre les différents systèmes a été conduit.

Ce travail mené par le ministère de la Santé et de la Prévention et l'AP-HP, en lien avec les éditeurs de systèmes d'information de laboratoire et Interop'Santé, a notamment conduit à des développements pour l'utilisation de normes d'échange (HL7 ; Hⁱ) et référentiels LOINC (*Logical Observation Identifiers Names & Codes*) qui s'appuient sur une terminologie de référence internationale pour la codification des résultats d'examen de biologie dans le but de standardiser les informations envoyées par les laboratoires et de faciliter l'exploitation des bases de données.

Ce travail d'adaptation des infrastructures a dû être effectué en miroir, et la synchronisation à la fois par l'AP-HP pour la centralisation des données et l'envoi vers les partenaires aval, et par Santé publique France pour la réception des données anonymisées, et ce à plusieurs reprises. Un important dispositif de monitoring des flux de données a dû être mis en place, afin de détecter et résoudre toute anomalie de transmission de manière très réactive.

Le recueil de données de santé individuelles impose un contexte de sécurité informatique et de protection des données très importantes. La mise en œuvre de la collecte et du traitement des données

⁽⁴⁾ <https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/consortium-emergen>

⁽⁵⁾ Santé publique France. Géodes – Taux d'incidence – Semaine glissante (pour 100 000 hab.) – 0-2 ans. https://geodes.santepubliquefrance.fr/#c=indicator&f=02&i=sp_7j_cage_scol.ti&s=2023-06-21-2023-06-27&t=a01&view=map2

s'est effectuée dans le respect des principes posés par le Règlement général sur la protection des données (RGPD) et de la loi informatique et libertés. Une homologation au référentiel général de sécurité, une certification hébergeur de données de santé, et des analyses d'impact relatives à la protection des données ont notamment été mises en place par l'ensemble des parties prenantes pour garantir la sécurité et la protection des données. L'AP-HP a renforcé la sécurisation de SI-DEP avec des outils de cyberattaque du fait de risques accrus liés à ce SI stratégique et national. La Commission nationale de l'informatique et des libertés (Cnil) a également diligencé de nombreux contrôles sur pièces et sur place du SI-DEP, et a rendu des avis favorables aux évolutions du décret afférent pris en conseil d'État.

Les enjeux scientifiques et de santé publique

La surveillance virologique a évolué tout au long de la crise, en accord avec les pratiques de test, l'évolution du SARS-CoV-2 et l'émergence de variants et mutations, et la mise en place de mesures de gestion.

L'émergence des variants fin 2020 a représenté un défi majeur. Le choix de la nomenclature à adopter (Pangolin, Nexstrain, GISAID – *Global Initiative on Sharing Avian Influenza Data* –, ...) a nécessité de nombreuses discussions entre le CNR-VIR, les partenaires, ainsi que les laboratoires et les éditeurs de systèmes d'information des laboratoires (SIL), impliquant des délais importants pour prise en compte dans SI-DEP. Les difficultés liées à une intégration tardive dans le système d'information, la complétude du champ manuel par les laboratoires et les règles de protection des données spécifiques en génétiques n'ont pas permis d'exploiter correctement cette information, en raison d'un manque d'exhaustivité et de qualité des données saisies. La mise en place du projet Emergen a permis de suivre l'évolution génétique du SARS-CoV-2 pour détecter l'émergence et la distribution spatio-temporelle de variants.

Le changement de la stratégie de criblage d'une recherche de variant à une recherche de mutation fin 2020 a été accompagnée par une nouvelle nomenclature à adopter, ce qui a impacté la remontée des informations dans SI-DEP. Cela a nécessité une nouvelle phase d'adaptation par les laboratoires pour assurer une remontée des informations à un niveau national.

Santé publique France a construit des indicateurs épidémiologiques d'une forte précision pour suivre au mieux l'épidémie. La production quotidienne d'indicateurs permettant le suivi de l'épidémie à des échelons géographiques infra-départementaux (communes, EPCI, Iris), et pour différentes classes d'âges, a représenté un enjeu majeur pour la continuité d'activité.

À la précision des indicateurs calculés pour décrire au mieux la population, s'ajoute l'enjeu d'une production quotidienne d'un grand nombre d'informations selon des formats différents (fichiers de données, bulletins épidémiologiques, tableaux de bord, etc.),

ce qui a nécessité la mise en place de solutions innovantes pour l'automatisation et le monitoring des traitements, ainsi qu'une très importante mobilisation humaine de l'ensemble des acteurs.

L'importance de ces données pour le suivi de la pandémie a imposé de fortes contraintes pour garantir l'opérationnalité du dispositif sans discontinuité, nécessitant par exemple la mise en place d'astreintes de soir et week-ends pour la mise en relation de l'ensemble des acteurs impliqués dans le dispositif, depuis la collecte des données, la centralisation et leur stockage, l'envoi et le traitement par les agences sanitaires. Des modes dégradés de transmission des données ont pu être développés pour pallier des anomalies ponctuelles.

Les enjeux de partenariat et de communication

La collaboration rapprochée entre les acteurs institutionnels et de terrain a été un enjeu majeur dans la mise en place, l'évolution et la maintenance d'un tel dispositif. Le développement et les évolutions successives du système ont nécessité une mobilisation de tous les instants de l'ensemble des acteurs impliqués, depuis les professionnels de santé des secteurs publics et privés du fait de leurs pratiques, les éditeurs de solution informatique, l'intégration des évolutions par l'AP-HP, jusqu'aux agences sanitaires pour le recueil des données et la restitution d'indicateurs épidémiologiques, le tout sous le pilotage du ministère de la Santé et de la Prévention.

Dans ses missions de mise à disposition d'indicateurs épidémiologiques, Santé publique France a été conduit à communiquer auprès d'un large public, dans l'objectif d'accompagner les évolutions du dispositif. De nombreuses notes méthodologiques ont été rédigées à destination des autorités sanitaires, des partenaires, de la presse, mais également auprès du grand public.

La publication de ces informations en *open data* a nécessité une pédagogie importante au regard des indicateurs produits, et devait être accompagnée par une communication détaillée pour être compréhensible et réutilisable par le grand public. Des séances de décryptages hebdomadaires ont été organisées par le service de la communication de Santé publique France pour répondre aux différentes interrogations des journalistes et présenter les évolutions du dispositif en toute transparence. Ces indicateurs ont pu être repris par de nombreuses personnes, professionnelles ou non, et ont permis l'alimentation de nombreux tableaux de bord et outils de visualisation (TousAntiCovid ou CovidTracker⁽⁶⁾ par exemple).

Conclusion

Les expériences internationales et leçons tirées

D'autres expériences similaires ont été observées à l'échelle européenne pour suivre l'évolution de la

⁽⁶⁾ <https://covidtracker.fr/>

pandémie. La flexibilité et l'adaptabilité des systèmes d'information pour la surveillance virologique ainsi que les réseaux préexistants entre les laboratoires, les professionnels de santé et les acteurs de santé publique ont facilité cette transition, et renforcé la coopération entre eux⁹. Face à cette crise, plusieurs pays européens ont adapté les systèmes de surveillance virologique déjà en place pour pouvoir suivre la propagation du virus à des niveaux très fins, collecter des données en temps réel, mettre en place des mesures de prévention et suivre leurs impacts⁹.

Tout comme la France, l'Allemagne a adapté son système de surveillance déjà existant (pour le suivi de la résistance aux antimicrobiens) pour recevoir les résultats de tests de SARS-CoV-2 effectués en laboratoires de manière automatique et suivre l'épidémie en temps réel. Ce système couvre l'ensemble des laboratoires sur le territoire qui traitent et analysent les échantillons provenant des hôpitaux, centres de soins primaires et services ambulatoires¹⁰. La Belgique a également mis en place un système de surveillance exhaustif comparable à la France qui a permis de récolter les données de tests de laboratoires quotidiennement, en passant dans un premier temps par une centralisation des résultats via le réseau de laboratoires sentinelles belges, suivie par la mise en place d'une base de donnée centralisée pour la remontée des résultats directement à partir des SI des laboratoires¹¹. L'Espagne a adapté son système de surveillance électronique (SiViES) pour inclure la surveillance du SARS-CoV-2⁹.

Si les dispositifs de collecte de données étaient comparables, les définitions de cas et les modalités de calcul des indicateurs épidémiologiques pouvaient différer d'un pays à l'autre, conduisant à des difficultés dans les comparatifs des tendances entre les pays. SI-DEP a un caractère exceptionnel dans son périmètre de fonctionnalité en comparaison aux autres systèmes européens, notamment au travers de fonctionnalités tournées vers les citoyens pour remettre des certificats numériques à tous les patients dans le cadre de la mise en place du passe sanitaire.

Le partage de ces expériences permet notamment aux autorités telles que l'ECDC de produire des recommandations sur les évolutions à développer et maintenir pour faire face à de futures pandémies, comme la mise en place de systèmes de surveillance intégrés et représentatifs pour le suivi des infections respiratoires¹².

Les perspectives de SI-DEP

La pandémie de Covid-19 a démontré l'importance d'avoir accès à des données de qualité de manière réactive pour pouvoir mettre en place des actions adaptées dans un contexte de crise sanitaire¹³. Les investissements effectués dans le champ du numérique en santé en France depuis plus de 10 ans comme la mise en place de la carte e-CPS, France

Connect, du système Sesam-Vitale, etc., et l'usage généralisé d'outils numériques via les téléphones ou les PC a facilité l'utilisation et l'exploitation de la donnée dans ce contexte particulier.

L'utilisation de formats d'échanges standardisés et des référentiels internationaux permettent d'accélérer la mise en place de ces systèmes, et d'optimiser les échanges et l'interopérabilité des systèmes à des échelles nationales et internationales.

La mise en place de systèmes de surveillance exhaustifs permettant de répondre à plusieurs objectifs a démontré la nécessité de pouvoir traiter une quantité massive d'informations et de disposer d'infrastructures techniques robustes et prêtes à collecter et absorber des données en grande quantité de manière réactive (en temps de crise ou non). Ce besoin a mis en lumière l'importance des métiers de l'informatique pour la mise en place, la maintenance et l'évolution de tels systèmes, et de la *data science* pour l'exploitation massive et automatisée des données.

La pandémie liée au SARS-CoV-2 a mis en avant le besoin en santé publique de s'appuyer sur des systèmes d'information nationaux et exhaustifs, notamment à partir des données de biologie médicale pour le suivi de nombreuses maladies. L'avenir des systèmes de surveillance repose ainsi notamment sur l'exploitation plus systématique des données de santé générées par les professionnels de santé et déjà disponibles au sein de bases de données structurées, et sur leur diffusion via des SI modernes et automatisés, permettant d'alléger la charge de travail des déclarants, tout en garantissant la sécurité des données des patients.

Anticiper la mise en place de tels dispositifs est capital, afin de renforcer les surveillances existantes et d'anticiper la survenue future d'événements pour lesquels il sera nécessaire de disposer de données de manière réactive et exhaustive.

Le système SI-DEP pourra servir d'expérience et de modèle dans le cadre de la mise en place de futurs systèmes de surveillance intégrés, étendus à d'autres pathologies, ne se limitant pas aux maladies infectieuses.

La publication du décret n° 2023-700 du 31 juillet 2023¹⁴ et de l'arrêté du 7 août 2023¹⁵ relatif à la transmission obligatoire de données individuelles à l'autorité sanitaire et à la création du traitement de données à caractère personnel « LABOé-SI » garantit la remontée des résultats des laboratoires publics et privés, et permettra la pérennisation de la surveillance virologique pour la Covid-19, mais également l'intégration progressive d'autres maladies. Ce système aura vocation à être utilisé à la fois pour les surveillances pérennes et pour la gestion de crise sanitaire exceptionnelle via un module de crise qui permettra la mobilisation d'autres professionnels de santé. ■

Remerciements

Les auteurs remercient Corinne Danan (venue en renfort de l'Anses à Santé publique France de mars à mai 2020) pour sa contribution aux travaux préparatoires à la mise en place

de l'outil SI-DEP, tous les acteurs publics et privés ayant contribué à la création de cet outil, et tous les laboratoires de biologie médicale et professionnels de santé ayant contribué à l'alimentation de cette base de données nationale.

Liens d'intérêt

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

[1] Figoni J, Campèse C, Spaccferri G, Rolland P, Caserio-Schönemann C, Che D. Structuration évolutive d'une surveillance multi-sources pour répondre à une infection émergente : l'expérience française face à la Covid-19. *Bull Épidémiol Hebd.* 2023;(1):2-16. http://beh.santepubliquefrance.fr/beh/2023/1/2023_1_1.html

[2] Ministère des Solidarités et de la Santé. Arrêté du 7 mars 2020 portant modification de la liste des actes et prestations mentionnée à l'article L. 162-1-7 du code de la sécurité sociale (inscription de la détection du génome du SARS-CoV-2 par RT PCR). *JORF.* 2020;(0058):207-8. <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000041698000>

[3] Loi n° 2020-290 du 23 mars 2020 d'urgence pour faire face à l'épidémie de covid-19 (1). *JORF.* 2020;(0072):15-26. <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000041746313>

[4] Ministère des Solidarités et de la Santé. Décret n° 2020-551 du 12 mai 2020 relatif aux systèmes d'information mentionnés à l'article 11 de la loi n° 2020-546 du 11 mai 2020 prorogeant l'état d'urgence sanitaire et complétant ses dispositions. *JORF.* 2020;(0117):54-9. <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000041869923/>

[5] Institut national de la statistique et des études économiques (Insee). IRIS – Définition. 2016. <https://www.insee.fr/fr/metadonnees/definition/c1523>

[6] Spaccferri G, Larrieu S, Pouey J, Calba C, Benet T, Sommen C, *et al.* Early assessment of the impact of mitigation measures to control COVID-19 in 22 French metropolitan areas, October to November 2020. *Euro Surveill.* 2020; 25(50):2001974. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.50.2001974>

[7] Organisation mondiale de la santé. Suivi des variants du SARS-CoV-2. 2024. <https://www.who.int/fr/activities/tracking-SARS-CoV-2-variants>

[8] Santé publique France. COVID-19 – Dépistage réactif dans les écoles primaires : résultats de l'évaluation. 2021. <https://www.santepubliquefrance.fr/les-actualites/2021/covid-19-de-pistage-reactif-dans-les-ecoles-primaires-resultats-de-l-evaluation>

[9] Bagaria J, Jansen T, Marques DF, Hooiveld M, McMenamin J, de Lusignan S, *et al.* Rapidly adapting primary care sentinel surveillance across seven countries in Europe for COVID-19 in the first half of 2020: strengths, challenges, and lessons learned. *Euro Surveill.* 2022;27(26):2100864. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2022.27.26.2100864>

[10] Rigoine de Fougerolles T, Puig-Barbera J, Kassianos G, Vanhems P, Schelling J, Crepey P, *et al.* A comparison of coronavirus disease 2019 and seasonal influenza surveillance in five European countries: France, Germany, Italy, Spain and the United Kingdom. *Influenza Other Respir Viruses.* 2022; 16(3):417-28.

[11] Meurisse M, Lajot A, Dupont Y, Lesenfants M, Klamer S, Rebolledo J, *et al.* One year of laboratory-based COVID-19 surveillance system in Belgium: main indicators and performance of the laboratories (March 2020–21). *Arch Public Health.* 2021;79(1):188.

[12] European Centre for Disease Prevention and Control, World Health Organization Regional Office for Europe. Operational considerations for respiratory virus surveillance in Europe. Stockholm: ECDC; 2022. 37 p. <https://www.ecdc.europa.eu/en/publications-data/operational-considerations-respiratory-virus-surveillance-europe>

[13] Organisation for Economic Co-operation and Development. Open data in action – Initiatives during the initial stage of the COVID-19 pandemic. Paris: OECD; 2021. 28 p. <https://web.archive.org/2021-12-17/549237-use-of-open-government-data-to-address-covid19-outbreak.htm>

[14] Ministère de la Santé et de la Prévention. Décret n° 2023-700 du 31 juillet 2023 relatif à la transmission obligatoire de données individuelles à l'autorité sanitaire et à la création du traitement de données à caractère personnel « LABOé-SI ». *JORF.* 2023;(0176):53-6. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000047911698>

[15] Ministère de la Santé et de la Prévention. Arrêté du 7 août 2023 relatif au système d'information « LABOé-SI » et pris en application des articles R. 3113-5 et R. 1413-58-1 du code de la santé publique. *JORF.* 2023;(0193):27. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000047984449>

Citer cet article

Durand J, Fayad M, Feri A, Truong E, Taisne B, Amy E, *et al.* Surveillance virologique en France : impact de la Covid-19 et perspectives. *Bull Épidémiol Hebd.* 2024;(20-21):440-9. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_1.html

LA SURVEILLANCE ÉPIDÉMIOLOGIQUE EN TEMPS PROCHE DU RÉEL DES HOSPITALISATIONS LIÉES À LA COVID-19 EN FRANCE, 2020-2023 : SI-VIC

// NEAR REAL-TIME EPIDEMIOLOGICAL SURVEILLANCE OF COVID-19-RELATED HOSPITAL ADMISSIONS IN FRANCE, 2020-2023: THE SI-VIC SURVEILLANCE SYSTEM

Yann Le Strat (yann.lestrat@santepubliquefrance.fr), Fanny Chereau

Santé publique France, Saint-Maurice

Soumis le 26.08.2024 // Date of submission: 08.26.2024

Mots-clés : Hospitalisation, Surveillance épidémiologique, Covid-19
// **Keywords:** Hospitalisation, Epidemiological surveillance, COVID-19

À la suite des attentats perpétrés à Paris en 2015, le ministère des Affaires sociales et de la Santé a mis en œuvre, après avoir obtenu l'autorisation de la Commission nationale de l'informatique et des libertés (Cnil) en juillet 2016, un traitement automatisé de données à caractère personnel ayant pour finalité l'établissement d'une liste unique des victimes d'attentats pour l'information de leurs proches par la cellule interministérielle d'aide aux victimes, sous l'acronyme SI-VIC (Système d'information pour le suivi des victimes)⁽¹⁾.

SI-VIC est déclenché lors d'attentats ou de situations sanitaires exceptionnelles. Il permet de renseigner et de mettre à jour les informations relatives à l'identité du patient pris en charge, la prise en charge (hospitalière ou médico-psychologique) du patient, la personne à contacter, proche du patient. Les objectifs de SI-VIC dans le cas d'une situation sanitaire exceptionnelle sont 1/ d'assurer l'identification et le dénombrement hospitalier, 2/ d'offrir une visibilité de l'impact de l'événement sur l'offre de soins (ventilation des patients dans les hôpitaux et leur gravité), et 3/ de faciliter l'accompagnement des victimes et de leurs proches. Il s'agit d'un outil administratif et de gestion, ne contenant pas de données médicales sur le patient.

Pour répondre à la situation sanitaire exceptionnelle engendrée par l'arrivée de l'épidémie de Covid-19, SI-VIC a été mobilisé dès mars 2020. En effet, il n'existait pas en 2020 en France de système de surveillance permettant de connaître, en temps proche du réel, le nombre de personnes hospitalisées par pathologie. Il n'était donc pas possible de mobiliser une source de données existante permettant de monitorer quotidiennement le nombre de personnes hospitalisées, leur localisation (hôpital, type de service) et quelques caractéristiques des patients hospitalisés (âge, sexe, etc.).

À partir du 1^{er} mars 2020, des professionnels de santé hospitaliers ont saisi des informations dans SI-VIC. Dès le 20 mars, Santé publique France a pu avoir accès à une partie de cette base de données, afin de construire des indicateurs épidémiologiques permettant de suivre quotidiennement et de manière réactive

le fardeau de la Covid-19 à l'hôpital. Jusqu'au 31 mars, les instructions aux établissements de santé pour créer un dossier patient dans SI-VIC étaient un diagnostic de Covid-19 confirmé biologiquement. À partir du 31 mars, les établissements de santé ont renseigné les patients hospitalisés présentant un diagnostic de Covid-19 confirmé biologiquement ou un scanner thoracique évocateur de diagnostic de Covid-19. Ces données ont été exploitées très vite par Santé publique France pour pouvoir construire et communiquer des premiers indicateurs épidémiologiques dans son point épidémiologique hebdomadaire dès le 24 mars¹. Puis, au cours de la première semaine d'avril 2020, Santé publique France a mis à disposition en open data ces indicateurs issus de SI-VIC⁽²⁾, à travers son site Géodes, ainsi que sur le site data.gouv⁽³⁾.

Les indicateurs épidémiologiques représentaient le nombre total de patients hospitalisés, déclinés par type de service (en réanimation ou soins intensifs, en soins de suite et de réadaptation (SSR) ou unités de soins de longue durée (USLD), en hospitalisation conventionnelle, autres types de service). Le nombre cumulé de personnes retournées à domicile était également calculé, ainsi que le nombre de personnes décédées. Ces indicateurs étaient déclinés par département et par sexe. Afin de contrôler le risque d'identification des patients, les indicateurs étaient également déclinés par classes d'âge décennales, mais uniquement à un niveau régional. Étaient distingués le nombre d'hospitalisations en cours et le nombre de nouvelles hospitalisations (incidence), afin de mieux appréhender la dynamique du fardeau à l'hôpital. Des indicateurs portaient spécifiquement sur les hospitalisations en soins critiques (services de réanimation ou soins intensifs), afin d'identifier la part des cas graves parmi les cas hospitalisés.

Ces indicateurs étaient déclinés par date de déclaration, c'est-à-dire par date de saisie par les professionnels

⁽²⁾ Santé publique France. Nombre de personnes actuellement hospitalisées avec diagnostic COVID-19. https://geodes.santepubliquefrance.fr/#c=indicateur&f=0&i=covid_hospit.hosp&s=2023-04-05&t=a01&view=map2

⁽³⁾ data.gouv.fr. Données hospitalières relatives à l'épidémie de COVID-19. <https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/>

⁽¹⁾ <https://esante.gouv.fr/SI-VIC>

de santé car ces dates étaient très bien renseignées (saisie automatique lors de la validation du dossier). Ces indicateurs ont ensuite été déclinés par date d'admission (saisie manuelle), permettant de se rapprocher davantage de la date d'infection par le SARS-CoV-2, mais ces indicateurs nécessitaient un délai de consolidation (déclaration en moyenne 3 jours après l'admission).

Dans son point épidémiologique hebdomadaire du 9 avril 2020², Santé publique France indiquait que depuis le 1^{er} mars 2020, 1 002 établissements avaient déclaré au moins une hospitalisation et recensait 57 021 patients ayant été hospitalisés, 7 091 patients étaient décédés (71% étaient âgés de 75 ans et plus) et 19 337 patients étaient retournés à domicile. L'âge médian des personnes ayant été hospitalisées depuis le 1^{er} mars et jusqu'au 7 avril était de 69 ans. Les régions Île-de-France, Grand Est et Auvergne-Rhône-Alpes rapportaient le plus grand nombre de patients admis en réanimation.

Les transferts entre établissements d'une même région étaient tracés, l'établissement de départ signalait le transfert dans l'application via la fonctionnalité « signaler un transfert » et l'établissement de destination le validait à l'arrivée du patient. En cas de transfert dans un établissement d'une autre région, il était demandé aux établissements de santé d'origine de cocher une case nommée « Evasan ».

À partir de janvier 2021, une nouvelle fonctionnalité de SI-VIC a permis à Santé publique France d'identifier la part des patients hospitalisés pour le traitement de la Covid-19, et ceux porteurs du SARS-CoV-2 mais hospitalisés pour d'autres motifs. En effet depuis le début de l'épidémie, les indicateurs

hospitaliers incluaient l'ensemble des patients positifs au SARS-CoV-2, que l'hospitalisation soit en lien ou non avec la Covid-19. Cela a permis de suivre quotidiennement la proportion de patients hospitalisés pour prise en charge de la Covid-19. Au cours de l'année 2021, la proportion de patients hospitalisés pour Covid-19 était de 87% parmi l'ensemble des hospitalisations déclarées dans SI-VIC. Cette distinction d'hospitalisations avec ou pour Covid-19 était essentielle, car les patients porteurs du SARS-CoV-2 mais hospitalisés pour d'autres motifs ne devaient pas être comptabilisés comme cas sévères, mais la prise en charge spécifique qui leur était nécessaire avait néanmoins un impact sur la tension hospitalière (isolement en chambre individuelle, mesures de protection des soignants, équipe de soignants dédiée voire réorganisation de services, etc.).

De nouveaux indicateurs ont ensuite été proposés, rapportés au nombre d'habitants : taux de nouvelles hospitalisations et nouvelles admissions en soins critiques sur 7 jours glissants, ainsi que les taux d'hospitalisation et d'hospitalisation en soins critiques en cours, disponibles selon trois modalités : hospitalisations avec infection SARS-CoV-2, hospitalisations pour Covid-19, hospitalisations pour autre motif avec infection SARS-CoV-2, la première modalité étant la somme des deux autres.

La figure 1 représente le taux d'incidence des hospitalisations pour 100 000 habitants, calculé sur 7 jours glissants, en distinguant les hospitalisations avec ou pour Covid-19, sur la période mars 2020-juin 2023. Elle représente également la part des hospitalisations en soins critiques. La figure 2 représente des taux d'hospitalisations par classes d'âges, cet

Figure 1

Taux d'incidence des hospitalisations (pour 100 000 habitants), calculé sur 7 jours glissants, France, du 1^{er} mars 2020 au 30 juin 2023

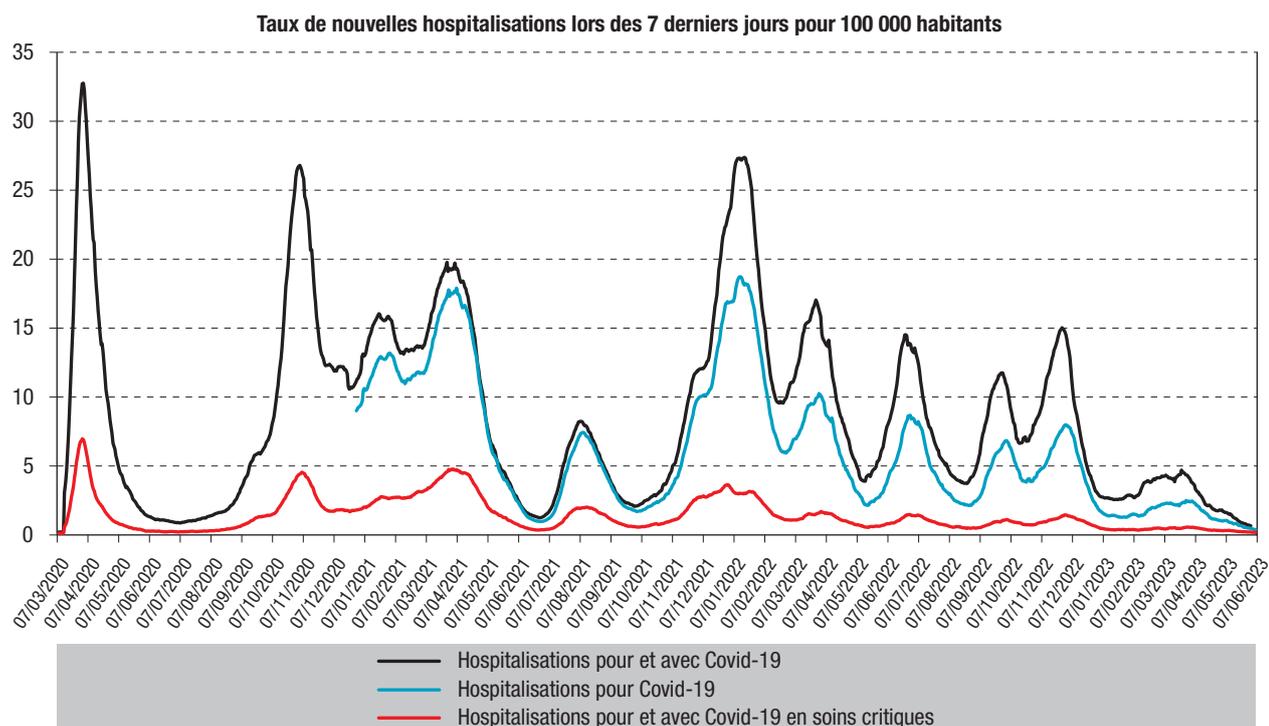


Figure 2

Taux d'incidence hebdomadaire des nouvelles hospitalisations pour Covid-19, par classe d'âge, France, du 1^{er} janvier 2021 au 30 juin 2023

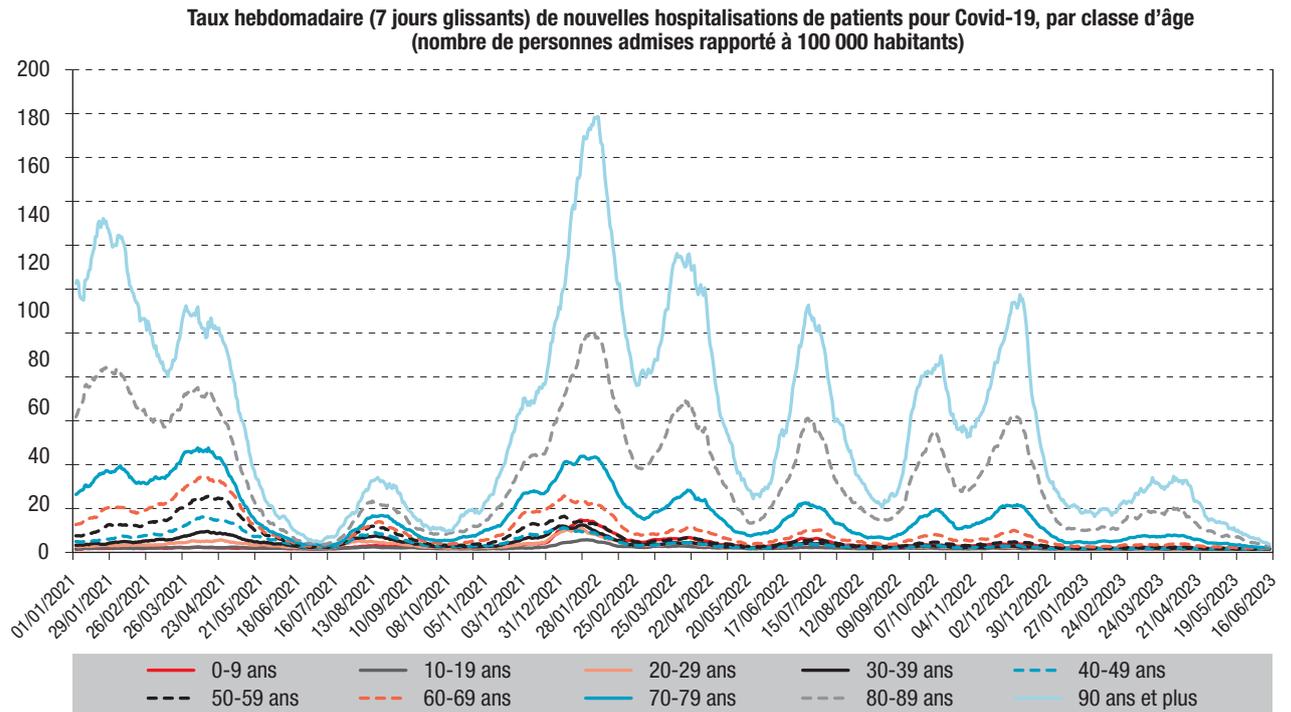
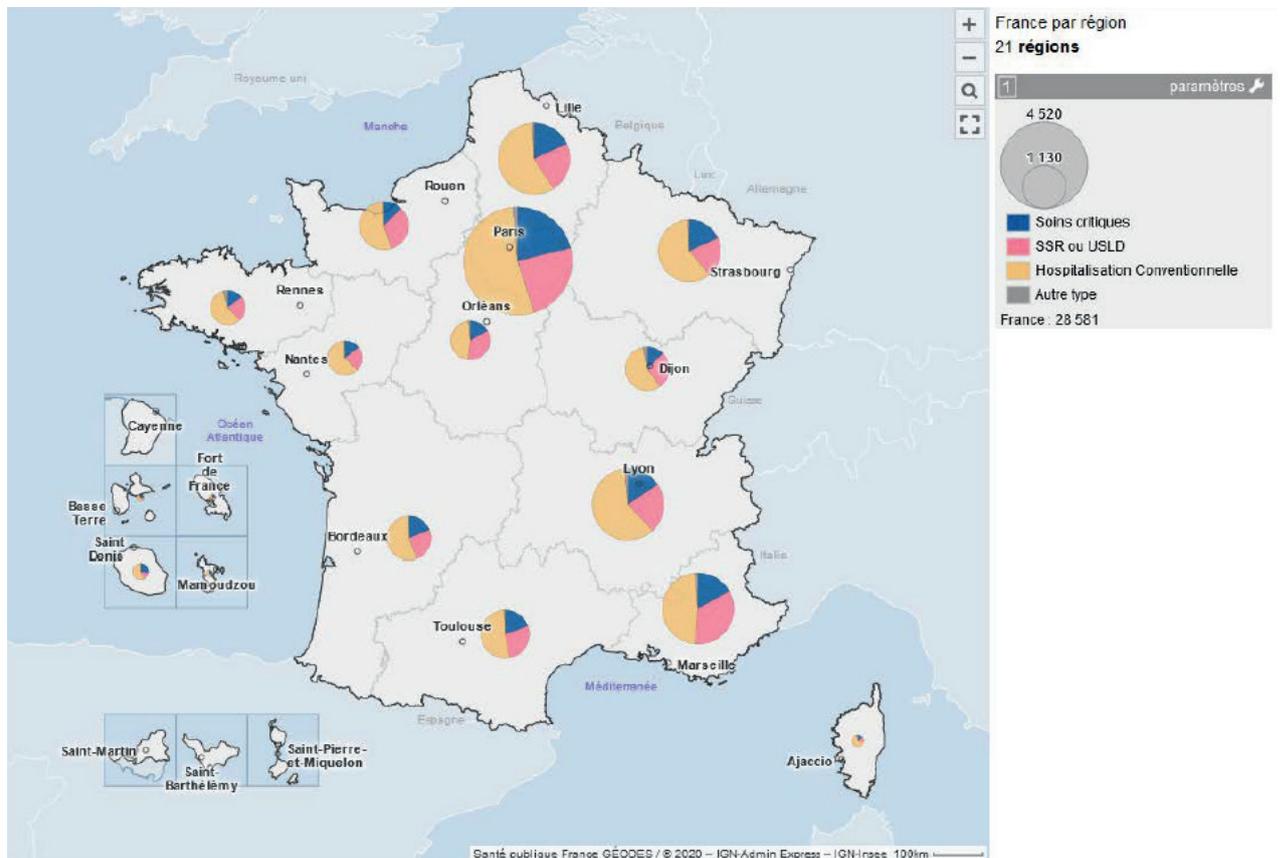


Figure 3

Nombre de personnes hospitalisées avec et pour Covid-19, par type d'hospitalisation par région, au 1^{er} avril 2021



SSR : soins de suite et de réadaptation ; USLD : unité de soins de longue durée.

Source : Santé publique France, Géodes.

https://geodes.santepubliquefrance.fr/#c=indicator&f=0&i=covid_hosp_type.type_hospit&s=2021-04-01&t=a01&view=map1

indicateur étant précieux pour monitorer les différentes dynamiques selon l'âge. Enfin la figure 3 représente le nombre de personnes hospitalisées avec et pour Covid-19, par type d'hospitalisation (soins critiques, SSR ou USLD, hospitalisations conventionnelles, autre type) et par région. Ces figures ne sont que trois illustrations de ce qui pouvait être construit à partir des données SI-VIC, mais illustrent les stratifications possibles des indicateurs en termes de temps, lieux, personnes.

En dehors du calcul d'indicateurs épidémiologiques, les données hospitalières ont été d'une grande utilité pour les travaux de modélisation. Couplées à d'autres données, les données hospitalières (incluant les données de mortalité à l'hôpital) issues de SI-VIC ont permis d'estimer très vite le fardeau de la Covid-19 en France^{3,4}, notamment l'impact du confinement et l'immunité de la population, ainsi que la proportion de la population ayant été infectée à un temps donné par le SARS-CoV-2 à un niveau régional.

Appariements avec les autres bases de données de la Covid-19

En 2021, les bases de données issues de SI-VIC, SI-DEP et VAC-SI, les systèmes d'information collectant de l'information respectivement sur les hospitalisations, les tests biologiques et la vaccination, ont pu être appariées selon un identifiant pseudonymisé⁵. Ces appariements ont été fondamentaux pour la réalisation d'analyses épidémiologiques plus poussées, puisqu'ils ont permis, entre autres, de mesurer la gravité selon le variant (notamment Omicron vs Delta), la présence de comorbidités et le statut vaccinal. Il a ainsi été montré que le risque d'événement grave était plus faible chez les cas Omicron que chez les cas Delta. Le risque d'événement grave augmentait avec la présence de comorbidités et chez les hommes, et était plus élevé chez les non-vaccinés que chez les primo-vaccinés⁶. Une autre étude a estimé que près de 480 000 hospitalisations, 132 000 admissions en unités de soins intensifs et 125 000 décès ont été directement évités par la vaccination chez les personnes âgées de 50 ans et plus entre décembre 2020 et mars 2022⁷. L'efficacité vaccinale et la durée de protection des vaccins Covid-19 à ARN ont également été estimées⁸.

Discussion

Le 30 juin 2023, la Direction générale de la santé a levé les consignes de saisie spécifiques dans SI-VIC pour les patients atteints de la Covid-19, étant donné l'amélioration de la situation sanitaire et l'impact limité sur l'offre de soins. Pendant plus de trois ans, l'application SI-VIC a été le moyen utilisé en France pour suivre l'épidémie de Covid-19 à l'hôpital, et pour compléter la surveillance multisource mise en place par Santé publique France⁹.

Elle a permis non seulement un monitoring quotidien de l'épidémie à l'hôpital sur tout le territoire français, mais a également nourri des modèles

mathématiques permettant de réaliser des projections de l'épidémie pour l'aide à la décision des autorités. Grâce aux données hospitalières appariées aux données des tests biologiques et de vaccination, il a été possible, par exemple, de démontrer l'efficacité de la vaccination pour réduire le nombre de formes graves.

Cependant cette surveillance hospitalière a reposé sur une application qui nécessitait de saisir des informations, pour chaque patient, par des professionnels de santé. Cette saisie a été chronophage et fastidieuse, à un moment où les professionnels de santé devaient avant tout prendre en charge les patients.

Afin de se préparer aux futures situations sanitaires exceptionnelles et notamment à l'émergence de nouveaux virus pouvant entraîner un fardeau important en termes de morbidité et/ou mortalité, la surveillance épidémiologique doit à l'avenir se moderniser, afin de pouvoir se reposer sur des données existantes sans devoir dépendre d'une saisie spécifique par les professionnels de santé. C'est le concept d'utilisation secondaire de données qui doit se généraliser, notamment à l'hôpital, pour que les données saisies par les professionnels de santé dans le cadre de la prise en charge de leurs patients puissent être extraites des bases de données hospitalières, afin de remplir des objectifs de surveillance épidémiologique.

Le projet nommé Orchidée (Organisation d'un réseau de centres hospitaliers impliqués dans la surveillance épidémiologique et la réponse aux émergences), coordonné par Santé publique France, a pour vocation de répondre à ce défi. Ce projet, financé en partie par la Commission européenne, qui a débuté en octobre 2024, implique 25 centres hospitaliers universitaires qui vont produire, à partir de leurs entrepôts de données de santé (EDS) des indicateurs épidémiologiques sur différentes thématiques, notamment infectieuses et en particulier des indicateurs concernant les infections respiratoires aiguës. ■

Liens d'intérêt

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

[1] Santé publique France. Covid-19 – Point épidémiologique du 24 mars 2020. Saint-Maurice: Santé publique France; 2020. 11 p. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-24-mars-2020>

[2] Santé publique France. Covid-19 – Point épidémiologique hebdomadaire du 09 avril 2020. Saint-Maurice: Santé publique France; 2020. 17 p. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-9-avril-2020>

[3] Salje H, Tran Kiem C, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, *et al.* Estimating the burden of SARS-CoV-2 in France. *Science*. 2020;369(6500):208-11.

[4] Hozé N, Paireau J, Lapidus N, Tran Kiem C, Salje H, Severi G, *et al.* Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: A modelling study. *Lancet Public Health.* 2021;6(6):e408-e15.

[5] Direction de la recherche, des études, de l'évaluation et des statistiques. Paris: Drees; 2021. Les appariements SI-VIC, SI-DEP et VAC-SI. <https://drees.solidarites-sante.gouv.fr/sources-outils-et-enquetes/les-appariements-si-vic-si-dep-et-vac-si>

[6] Auvigne V, Vaux S, Le Strat Y, Schaeffer J, Fournier L, Tamandjou C, *et al.* Severe hospital events following symptomatic infection with Sars-CoV-2 Omicron and Delta variants in France, December 2021-January 2022: A retrospective, population-based, matched cohort study. *EClinicalMedicine.* 2022;48:101455.

[7] Tan-Lhernould L, Tamandjou C, Deschamps G, Platon J, Sommen C, Chereau F, *et al.* Impact of vaccination against severe COVID-19 in the French population aged 50 years and above: A retrospective population-based study. *BMC Med.* 2023;21(1):426.

[8] Tamandjou Tchuem CR, Auvigne V, Vaux S, Montagnat C, Paireau J, Monnier Besnard S, *et al.* Vaccine effectiveness and duration of protection of COVID-19 mRNA vaccines against Delta and Omicron BA.1 symptomatic and severe COVID-19 outcomes in adults aged 50 years and over in France. *Vaccine.* 2023;41(13):2280-8.

[9] Figoni J, Campèse C, Spaccaferri G, Rolland P, Caserio-Schönemann C, Che D, *et al.* Structuration évolutive d'une surveillance multi-sources pour répondre à une infection émergente : l'expérience française face à la Covid-19. *Bull Épidémiol Hebd.* 2023;(1):2-16. http://beh.santepubliquefrance.fr/beh/2023/1/2023_1_1.html

Citer cet article

Le Strat Y, Chereau F. Focus. La surveillance épidémiologique en temps proche du réel des hospitalisations liées au Covid-19 en France, 2020-2023 : SI-VIC. *Bull Épidémiol Hebd.* 2024;(20-21):450-4. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_2.html

> ARTICLE // Article

VAC-SI : UN SYSTÈME D'INFORMATION POUR LE SUIVI DE LA COUVERTURE VACCINALE DES VACCINS CONTRE LA COVID-19

// VAC-SI: A SURVEILLANCE SYSTEM TO MONITOR COVID-19 VACCINATION COVERAGE

Johnny Platon¹ (johnny.platon@santepubliquefrance.fr), Laure Fonteneau¹, Rémi Hanguehard¹, Bertrand Gagnière¹, Gaëlle Gault¹, Guilhem Deschamps¹, Jérôme Brocca², Jean-François Girod³, Philippe Jouzier³, Sophie Vaux¹

¹ Santé publique France, Saint-Maurice

² Ministère de la Santé et de la Prévention, Paris

³ Caisse nationale de l'assurance maladie, Paris

Soumis le 02.08.2023 // Date of submission: 08.02.2023

Résumé // Abstract

Dans un contexte de pandémie, mesurer l'évolution et l'impact de la vaccination contre une maladie est essentiel pour la surveillance épidémiologique. Vaccin Covid est l'outil numérique permettant la traçabilité des vaccinations, et VAC-SI est un système d'information issu de Vaccin Covid, constituant la base de données des vaccinations contre la Covid-19 et des indicateurs de suivi dans le cadre de la pandémie.

Cet article décrit la mise en œuvre du système VAC-SI en pleine crise sanitaire, de la construction de la base de données avec les différents acteurs impliqués jusqu'à la production d'indicateurs, en passant par l'exploitation de la base par Santé publique France et ses évolutions.

Dans l'élaboration de ce nouveau système, de multiples défis techniques sont apparus face aux différents enjeux et besoins, ainsi qu'aux évolutions de la politique vaccinale, qui ont nécessité une capacité d'adaptation constante et une forte réactivité.

In the context of a pandemic, measuring the progression and impact of vaccination against a disease is essential for epidemiological surveillance. Vaccin Covid is a vaccination tracking tool and VAC-SI is an information system based on this tool. Together, they constitute France's COVID-19 vaccination database, which was used to produce indicators for monitoring the pandemic.

This article describes the implementation of the VAC-SI system during the health crisis, from the construction of the database with various stakeholders, to the production of indicators, including how Santé publique France used and adapted the system over time.

In developing this new system, multiple technical challenges arose in relation to different concerns and needs, along with changes in vaccination policy, which required a capacity for constant adjustment and a high level of responsiveness.

Mots-clés : Covid-19, Surveillance, Vaccination, Système d'information, Épidémiologie

// **Keywords:** COVID-19, Surveillance, Vaccination, Information systems, Epidemiology

Introduction

Le virus SARS-CoV-2 a émergé fin janvier 2020 en France. Les premières injections de vaccin contre la maladie émergente Covid-19 ont été administrées fin décembre 2020, soit un peu moins d'un an après les premiers cas d'infection en France. Les premiers vaccins mis à disposition en France étaient deux vaccins à ARNm : Cominarty®, développé par Pfizer-BioNTech, et Spikevax®, développé par Moderna. Deux vaccins à vecteur viral ont ensuite été rendus disponibles en France : Vaxzevria® d'AstraZeneca, puis Janssen® de Johnson & Johnson. Les vaccins contre la Covid-19 ont été dans un premier temps recommandés aux personnes les plus à risque de développer des formes graves de Covid-19 telles que les personnes âgées, les personnes résidant en collectivités de personnes âgées, les personnes atteintes d'une maladie chronique ainsi que les personnes à risque d'exposition tels que les professionnels de santé. Au fur et à mesure, la campagne a été étendue aux personnes plus jeunes jusqu'à couvrir l'intégralité de la population française à partir de janvier 2022, avec l'ouverture de la vaccination aux enfants.

Dès son introduction dans les établissements d'hébergement pour personnes âgées dépendantes (Ehpad), la vaccination contre la Covid-19 a eu un impact majeur sur les hospitalisations et sur la mortalité hospitalière¹. En diminuant par ailleurs les formes symptomatiques, elle a contribué à limiter la propagation du virus SARS-CoV-2¹⁻³. L'estimation régulière de la couverture vaccinale (CV) était indispensable pour suivre le déploiement de la campagne de vaccination. Des indicateurs de CV dans les populations pour lesquelles la vaccination était recommandée ont été produits et mis à disposition des autorités et de la population française par Santé publique France à partir du 28 janvier 2021. Ces indicateurs ont dû tenir compte des très nombreuses évolutions des recommandations de vaccination portant notamment sur les populations ciblées, les types de vaccins, l'évolution des schémas vaccinaux, le nombre de doses en lien avec les recommandations de la Haute Autorité de santé (HAS). Les estimations de CV contre la Covid-19 ont permis d'alimenter des modèles de progression de l'épidémie selon différents scénarios et de réaliser des études d'impact¹⁻³.

Afin de tracer les injections de vaccins contre la Covid-19, l'outil Vaccin Covid a été créé par la Caisse nationale d'assurance maladie (Cnam) dans des délais très courts. La base de données VAC-SI issue de Vaccin Covid a ensuite été créée et mise à disposition afin de produire les indicateurs dédiés au suivi de la vaccination. Tous ces composants constituent le système de surveillance VAC-SI. Santé publique France a mis en place une exploitation propre de la base VAC-SI, en collaboration avec ses partenaires, la Cnam et la Direction du numérique (DNUM), visant à répondre à la demande des autorités de disposer d'indicateurs pour la vaccination contre la Covid-19 (nombre de personnes vaccinées et CV) avec une très grande réactivité (données quotidiennes 7j/7),

pour les différentes populations et pour différents niveaux géographiques (national, régional, départemental). L'appariement des données du système de suivi des vaccinations (VAC-SI) et des données des tests biologiques effectués pour la recherche du SARS-CoV-2 (SI-DEP : Système d'information de dépistage populationnel) a également permis la réalisation d'estimations des efficacités vaccinales^{4,5}. À l'image des autres systèmes de surveillance liés au virus, le système VAC-SI devait aussi permettre une forte réactivité face à la propagation du virus et à ses mutations (variants). Le but de cet article est de décrire le processus de construction du système de surveillance, de Vaccin Covid jusqu'à la production d'indicateurs, en passant par la base de données.

La base de données VAC-SI

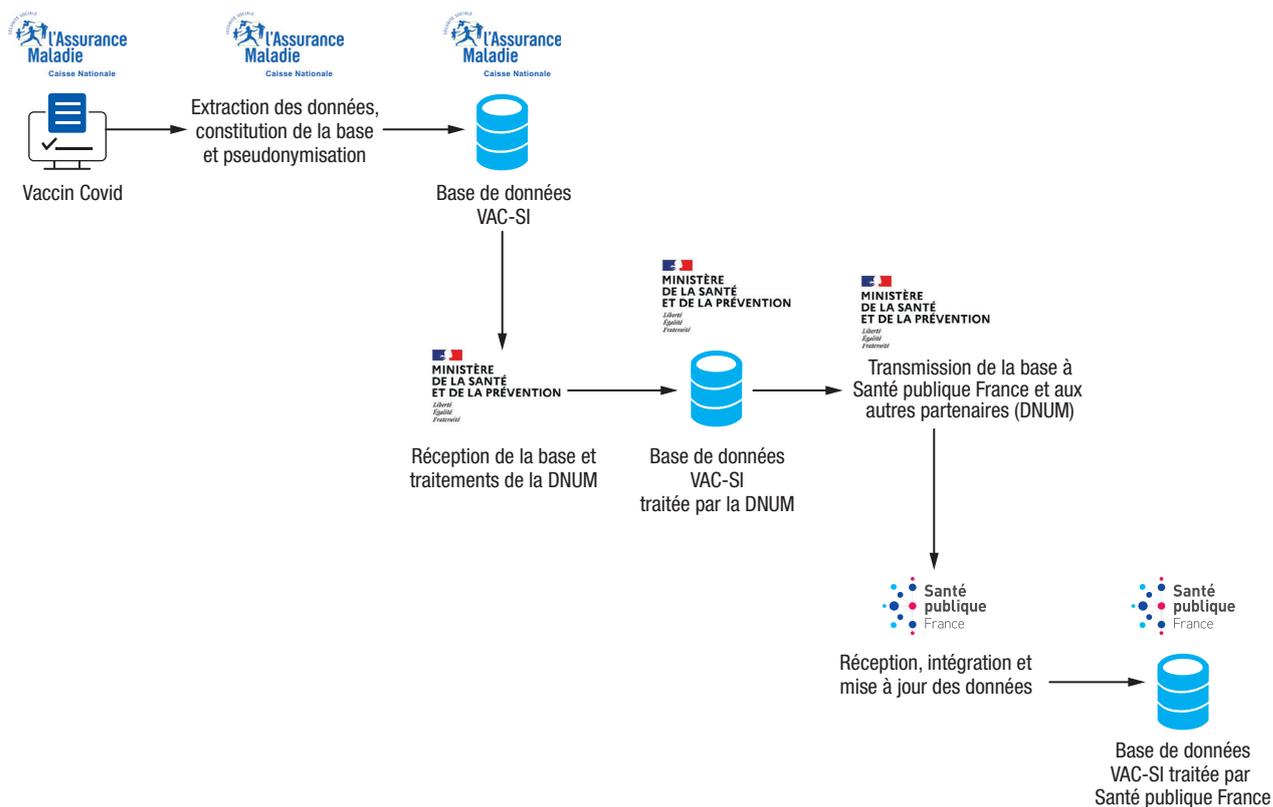
Le 4 janvier 2021, la Cnam a mis en service Vaccin Covid, outil pour le recueil spécifique des données de vaccination contre la Covid-19 et la traçabilité. Les professionnels vaccinateurs devaient y saisir les informations relatives à chaque injection. Une base de données, la base VAC-SI, a ensuite été créée afin de mettre à disposition les données issues de Vaccin Covid. Cette base contient ainsi l'ensemble des injections de vaccins contre la Covid-19 administrées en France et des informations utiles pour l'exploitation par l'ensemble des organismes destinataires de la base portant sur :

- les injections : date de l'injection, date de saisie dans Vaccin Covid, le rang vaccinal (première, deuxième, troisième dose), le motif de l'injection (primovaccination ou rappel), le nom technique du vaccin, la voie d'administration, le code postal du lieu de vaccination ;
- les patients : âge, sexe, année de naissance, et un pseudonyme ;
- les établissements de santé : catégorie du lieu de vaccination ;
- les professionnels de santé ayant réalisé l'injection : catégorie du professionnel de santé, spécialité ;
- les vaccins : nom commercial du vaccin ;
- des données additionnelles collectées par la Cnam sans être mises à disposition des professionnels de santé. Il peut s'agir de données relatives à la santé de la personne comme les critères médicaux ou non médicaux d'éligibilité à la vaccination. Ces données ont été utiles pour l'organisation de la campagne de vaccination, l'envoi des invitations aux personnes éligibles à la vaccination, la création de cohortes de sous-populations spécifiques.

Plusieurs partenaires sont intervenus dans la constitution et la gestion de la base VAC-SI (figure 1). La Cnam a mis à disposition quotidiennement la base à la DNUM pour une première exploitation de ces données. Afin de répondre aux exigences de la Commission nationale de l'informatique et des libertés (Cnil) et de la protection des données personnelles,

Figure 1

Description du processus de construction de la base VAC-SI



DNUM : Direction du numérique du ministère de la Santé et de la Prévention.

les données transmises étaient pseudonymisées par la Cnam. La DNUM procédait à des traitements dédiés à la qualité des données et à des ajouts de variables complémentaires (calcul du code département d'injection à partir des données de localisation de l'injection). Ces traitements et algorithmes de calcul ont été définis lors de réunions pluri-hebdomadaires et hebdomadaires d'un groupe de travail rassemblant notamment des intervenants de la Cnam, de la DNUM, de la Direction de la recherche, des études, de l'évaluation et des statistiques (Drees), des agences régionales de santé (ARS) et de Santé publique France.

L'objectif était de centraliser une partie des traitements sur les données afin d'homogénéiser les indicateurs produits pour les différents partenaires. À l'issue de ces premiers traitements, la DNUM transmettait quotidiennement la base de données aux différents partenaires (Santé publique France, Drees). La Drees exploite les données pseudonymisées issues des trois principales bases de données sur la crise Covid-19 : SI-VIC (hospitalisation), SI-DEP (dépistage) et VAC-SI (vaccination). Elle réalise des croisements entre ces bases, notamment un suivi selon le statut vaccinal des personnes testées positives au Covid-19 et des personnes hospitalisées. Santé publique France exploitait quotidiennement (7j/7) la base ainsi transmise, avec l'ajout de nouveaux tests de qualité et de données complémentaires pour ses propres analyses (données de population de l'Institut national de la statistique et des études économiques – Insee – par exemple). Le département

de résidence de la personne vaccinée a ainsi pu être ajouté à compter du 28/10/2021 grâce à une table de correspondance transmise par la Cnam à Santé publique France.

Production des indicateurs de couverture vaccinale

Les enjeux, besoins et résultats

Les indicateurs dédiés au suivi de la vaccination étaient les nombres de personnes vaccinées, ainsi que la CV qui correspond au nombre total de personnes vaccinées, divisé par la population ciblée. La couverture vaccinale était calculée pour plusieurs niveaux géographiques (nationale, régionale, départementale), par classes d'âge (5-9 ans, 10-11, 12-17, 18-24, 25-29, 30-39, 40-49, 50-59, 60-64, 65-69, 70-74, 75-79, et 80 ans et plus), et par sexe. Le nombre de personnes vaccinées par type de vaccin était également produit. Les nombres de personnes vaccinées et les couvertures vaccinales étaient calculés en population générale, ainsi que par type de population cible correspondant aux populations particulièrement à risque, ou particulièrement exposées à la Covid-19 et retenues comme populations prioritaires pour la vaccination par la HAS⁶ :

- les résidents en Ehpad ou en unités de soins de longue durée (USLD) ;
- les professionnels exerçant en Ehpad ou USLD ;
- les professionnels de santé.

Ces indicateurs de CV étaient produits en distinguant les personnes :

- vaccinées par au moins une dose ;
- ayant reçu une primovaccination complète correspondant aux personnes : vaccinées par deux doses pour les vaccins nécessitant deux doses (vaccins Pfizer, Moderna ou AstraZeneca) ; vaccinées par une dose par les vaccins nécessitant une seule dose (vaccin Janssen) ; vaccinées par une dose par les vaccins Pfizer, Moderna ou Astra-Zeneca en cas d'antécédent de Covid-19 ; vaccinées par trois doses pour les personnes particulièrement à risque, comme par exemple les personnes immunodéprimées ;
- ayant reçu une première dose de rappel ;
- ayant reçu une deuxième dose de rappel ;
- ayant reçu une troisième dose de rappel ;
- ayant reçu une dose de rappel avec un vaccin adapté au variant Omicron.

La figure 2 présente l'évolution de la vaccination par au moins une dose, primo-vaccination complète et première dose de rappel, pour la France entière, de fin décembre 2020 à juillet 2023.

Au total, environ quatre-vingts indicateurs de couverture vaccinale ont ainsi été produits quotidiennement depuis février 2021⁷. L'évolution des recommandations de vaccination a contraint à revoir à plusieurs reprises les populations suivies en termes notamment d'âge, de profil et de schéma vaccinal.

Ces indicateurs étaient mis à disposition pour les bilans quotidiens transmis au ministère de la Santé et publiés dans les points épidémiologiques, en open

data sur le site Géodes⁽¹⁾, les plateformes data.gouv⁽²⁾ et TousAntiCovid⁽³⁾, ainsi que sur le tableau de bord de Santé publique France⁽⁴⁾. Ces indicateurs étaient également transmis à partir de début octobre 2021 au *European Centre for Disease Prevention and Control* (ECDC), à la fréquence de deux fois par semaine puis une fois par semaine à partir de l'été 2022. Plusieurs pays de l'Union européenne ont pu collecter des données via un système de suivi de la vaccination et construire des indicateurs similaires. La figure 3, issue de l'article de van Kessel et coll.⁸, recense les indicateurs calculés pour 10 pays dont la France. Tous ces pays ont construit des indicateurs en fonction d'un premier schéma vaccinal suivi de doses de rappel, stratifiés par classes d'âge et de façon quotidienne. La France est l'un des seuls pays à avoir publié des indicateurs sur des sous-catégories de population comme les personnes vulnérables, comprenant les personnes à très haut risque de forme grave ou de décès, pouvant bénéficier d'affections de longue durée (ALD), les personnes avec transplantation d'organe solide, ou les personnes atteintes de certaines maladies rares⁽⁵⁾, de trisomie 21.

(1) <https://geodes.santepubliquefrance.fr/#c=indicateur&view=map2>

(2) <https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-personnes-vaccinees-contre-la-covid-19-1/>

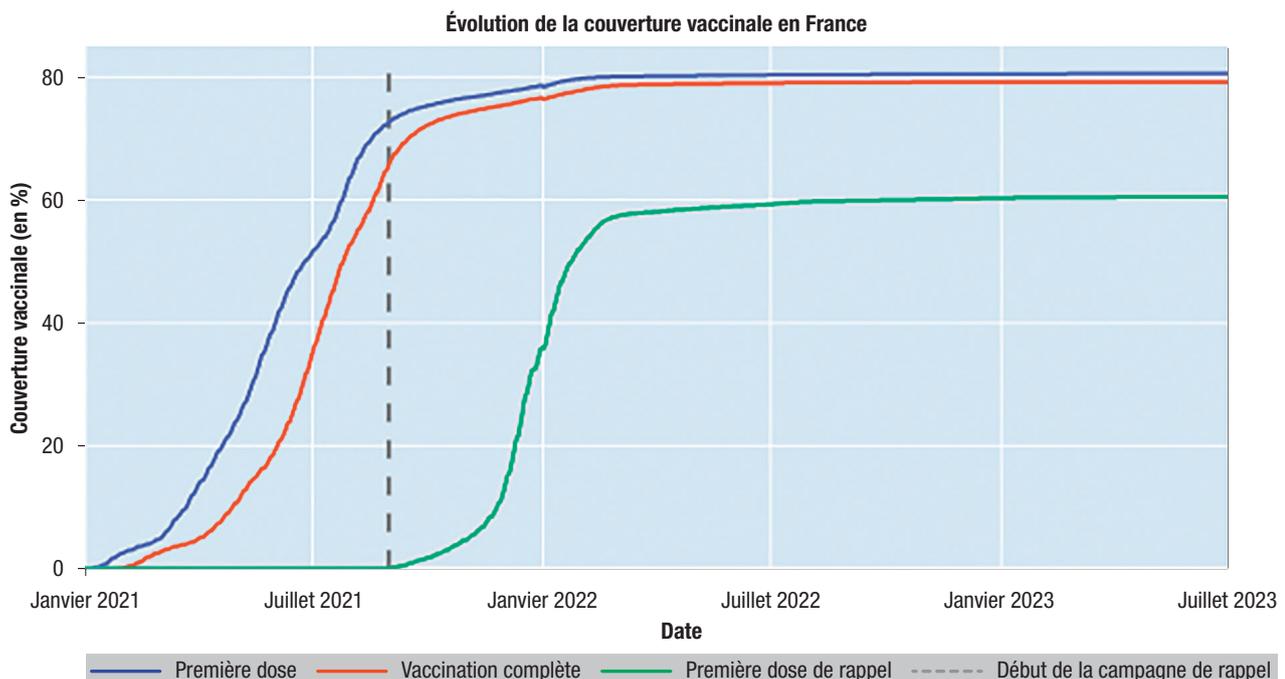
(3) <https://sante.gouv.fr/soins-et-maladies/maladies/maladies-et-infections-respiratoires/covid-19/article/tousanticovid>

(4) <https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolution-de-la-covid-19-en-france-et-dans-le-monde>

(5) Ministère du Travail, de la Santé et des Solidarités. La liste des comorbidités et des maladies associées à un risque de forme grave de Covid-19. 2023. <https://sante.gouv.fr/soins-et-maladies/maladies/maladies-et-infections-respiratoires/coronavirus-11069/vaccin-covid-19/article/la-liste-des-comorbidites-et-des-maladies-associees-a-un-risque-de-forme-grave>

Figure 2

Évolution de la vaccination contre la Covid-19 en France, de fin décembre 2020 à juillet 2023



Enrichissement et évolution de la base de données

Lieu de résidence

Du fait de l'absence de transmission du lieu de résidence des personnes vaccinées dans la base VAC-SI, et de la pseudonymisation des données transmises, Santé publique France était dans l'impossibilité de produire des données de CV par lieu de résidence. Dans un premier temps, les données de CV ont ainsi, par défaut, tenu compte du département de vaccination comme approximation du département de résidence. L'estimation de la CV s'appuyant sur l'estimation de la population transmise par l'Insee, les estimations de CV départementales pouvaient ainsi être biaisées lorsque des personnes se faisaient vacciner dans un autre département que leur département de résidence. Une correction a pu être apportée grâce à la transmission par la Cnam à Santé publique France à partir du 28 octobre 2021 d'une table de correspondance permettant de lier les

personnes vaccinées à leur département de résidence tout en préservant la confidentialité. Une mise à jour des départements de résidence a pu être réalisée en novembre 2022 (figure 4).

Décès

La base de données VAC-SI contient l'ensemble des injections des patients y compris des patients décédés depuis le début de la campagne de vaccination. L'information du décès n'était pas disponible dans VAC-SI. La taille de la population prise en compte pour le dénominateur de la CV (données Insee) prend en revanche bien en compte les décès (et les naissances survenues). La conservation des personnes vaccinées décédées a ainsi pu conduire à sur-estimer la CV, notamment dans les classes d'âge les plus avancées. La date de décès des personnes vaccinées a pu être ajoutée à la base à partir du 13 octobre 2022 grâce à une autre table de correspondance transmise par la Cnam.

Figure 3

Disponibilité des indicateurs publiés liés à la vaccination contre la Covid-19 dans 10 pays européens

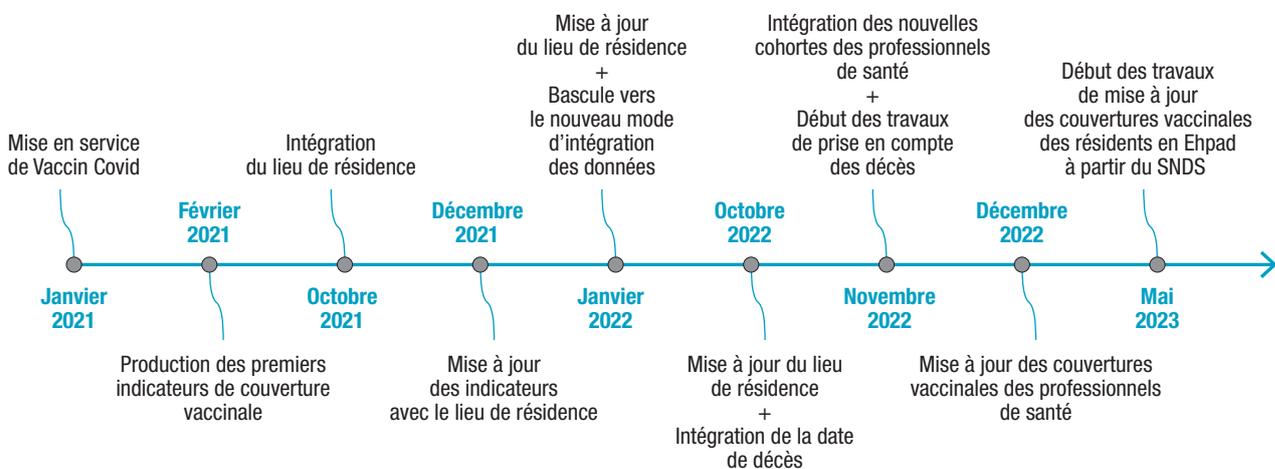
	Allemagne	Angleterre	Australie	Danemark	Espagne	France	Irlande	Italie	Pays-Bas	Pologne
Nombre total de personnes vaccinées	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Nombre de personnes vaccinées par une dose	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Nombre de personnes vaccinées par deux doses	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Nombre de personnes vaccinées avec un rappel	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Nombre de personnes vaccinées avec deux rappels	Oui	Oui	Non renseigné	Oui	Oui	Oui	Oui	Oui	Oui	Non renseigné
Nombre de personnes vaccinées par type de vaccin	Oui	Non	Non	Non	Oui	Oui	Oui	Oui	Oui	Non
Nombre de personnes vaccinées par classes d'âge	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Nombre de personnes vaccinées par régions	Oui	Oui	Oui	Oui	Oui	Oui	Non	Oui	Oui	Oui
Nombre de personnes vaccinées par type de professionnel vaccinateur	Oui	Non	Non	Non	Non	Oui	Non	Oui	Non	Non
Nombre de personnes vaccinées par jour	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Nombre de personnes vaccinées par sexe	Non	Non	Oui	Oui	Oui	Oui	Non	Oui	Oui	Oui
Vaccinations des personnes vulnérables	Non	Non	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui

■ Oui
 ■ Non renseigné
 ■ Non

Source : van Kessel et coll. [8]

Figure 4

Évolution de la base de données VAC-SI dans le temps



Ehpad : établissements d'hébergement pour personnes âgées dépendantes ; SNDS : Système national des données de santé.

Professionnels de santé

Du fait de la pseudonymisation des données dans la base VAC-SI, Santé publique France n'était pas en mesure d'identifier les personnes dans VAC-SI qui étaient des professionnels de santé. Au lancement de la campagne de vaccination, la Cnam pouvait techniquement repérer certains professionnels de santé, renseignés dans la base VAC-SI, essentiellement des libéraux ainsi que certains professionnels exerçant dans les établissements de santé ou des établissements médico-sociaux identifiés grâce à une saisie manuelle rendue possible dans Vaccin Covid. Cette identification de professionnels était cependant non exhaustive et ne permettait pas d'identifier les professionnels de santé salariés. Des imprécisions pouvaient également survenir en cas d'erreur lors de la saisie manuelle. Avec le temps, la qualité de la cohorte de professionnels ainsi constituée a décliné du fait de l'impossibilité de la prise en compte des décès, des embauches ou des licenciements. Ce n'est qu'à partir de novembre 2022, en s'appuyant sur le Répertoire partagé des professionnels de santé (RPPS) et le répertoire Adeli (automatisation des listes) que la Cnam et la Drees ont pu fournir à Santé publique France une nouvelle table de correspondance, résultat de l'appariement entre les répertoires et Vaccin Covid, qui a été intégrée en base à partir du 23 novembre 2022 et a permis d'obtenir des cohortes de professionnels de santé de meilleure qualité.

Résidents en Ehpad

En raison de l'absence d'information dans la base VAC-SI pour identifier les personnes résidentes en Ehpad, population particulièrement à risque pour la Covid-19, nous nous sommes initialement appuyés sur l'âge (à partir de 65 ans) et le lieu de vaccination (Ehpad) pour tenter d'identifier ces résidents. Les couvertures vaccinales ainsi produites ont été cependant surestimées du fait de la vaccination dans ces établissements, notamment en début de campagne, de personnes qui n'étaient pas des résidents et ce pour éviter les pertes de doses (conditionnements multiples). La base Resid-ESMS du Système national des données de santé (SNDS) constitue un référentiel des résidents en Ehpad et USLD (unité de soins de longue durée), ainsi que des bénéficiaires d'ESMS (Établissement social ou médico-social). Grâce à l'exploitation de la base Resid-ESMS, la Cnam a pu fournir à Santé publique France, une table de correspondance permettant d'identifier les résidents en Ehpad et permettant de produire des estimations de meilleure qualité. La non mise à jour régulière de cette cohorte peut cependant conduire à des défauts de qualité des estimations du fait de la non prise en compte des nouvelles admissions ou des décès par exemple. À partir de juin 2023, l'estimation des CV chez les résidents en Ehpad a pu être produite directement par Santé publique France via les données du SNDS, grâce au croisement de la base Resid-ESMS et d'une nouvelle base intégrée au SNDS qui recense les vaccinations contre la Covid-19 issues de Vaccin Covid.

Ressources informatiques et traitements des données

L'ajout des différentes données externes précédemment listées sont venues complexifier l'exploitation de la base VAC-SI au cours du temps. Ces adaptations ont dû être réalisées manuellement et sans automatisation possible lors de la mise à disposition ponctuelle des tables de correspondance. L'acquisition de ces données a également été soumise à un circuit de validation juridique, visant à attester que la transmission et l'utilisation des données s'inscrivaient dans le cadre d'obligations du code de la santé publique et dans le respect du Règlement général sur la protection des données (RGPD). Chacune de ces évolutions a conduit à la mise en place de nouveaux tests d'assurance qualité et au besoin d'accompagnement pour expliquer les possibles modifications de couvertures vaccinales qu'elles induisaient.

La base est devenue très rapidement volumineuse et les traitements de plus en plus coûteux en ressources informatiques et en temps de calcul (jusqu'à 4h de traitement de la réception des données jusqu'à la production des indicateurs pour une base ayant atteint, à partir de début octobre 2021, environ 100 millions d'injections). Afin d'éviter les retards dans les productions quotidiennes d'indicateurs et dans un souci de réactivité du système, une solution a été mise en place fin janvier 2022. Initialement, les données reçues la veille étaient entièrement remplacées par les nouvelles données du jour. La solution a consisté à conserver les données historiques et d'ajouter quotidiennement dans la base uniquement les nouvelles injections et d'apporter les modifications et corrections ou suppressions concernant des injections antérieures. Dans ce nouveau processus, les données de la veille devaient ainsi être appariées avec les données du jour pour procéder aux modifications. Ce nouveau mode de mise à jour de la base a nécessité des évolutions sur toute la chaîne depuis la constitution de la base par la Cnam, faisant intervenir de nouveau l'ensemble des partenaires de la chaîne.

Un maintien des deux modes d'intégration des données a été nécessaire durant quelques mois, afin de s'assurer du bon fonctionnement.

Discussion

Les points forts

La mise en place du système VAC-SI a été très rapide, constituant un nouveau dispositif exhaustif du suivi de la vaccination à grand échelle. Il permettait de disposer d'un système réactif de l'adhésion à une campagne de vaccination, avec une remontée très rapide des données à J-1.

L'exploitation de la base de données par Santé publique France a également permis une mise à jour quotidienne des estimations de CV, ce qui a apporté des informations quotidiennement aux autorités

sanitaires, aux décideurs et au grand public pendant la crise par le biais des bilans quotidiens, points épidémiologiques et via l'open data.

Le travail collaboratif entre les différentes entités Cnam, DNUM, Drees, ARS et Santé publique France a été très bénéfique. Les réunions techniques avec ces différents organismes pour fixer les indicateurs, les modes de calcul, les algorithmes et les contrôles qualité, ont permis d'homogénéiser l'exploitation de la base et d'avoir une forte réactivité sur les évolutions survenues dans la base. Grâce à l'implication très forte des différents partenaires, il a été possible de produire pour la première fois et à l'échelle de toute une population, des couvertures vaccinales quasiment en temps réel, c'est-à-dire avec prise en compte des vaccinations ayant été réalisées la veille.

L'exploitation de la base a été conçue dès le départ autour de traitements automatisés et paramétrés permettant de répondre très vite à la fréquence souhaitée de production des indicateurs et aux évolutions. De solides procédures de contrôles qualité ont été mises en place en base de données et sur les résultats des calculs des couvertures vaccinales, limitant les risques d'incohérences sur les indicateurs produits par les différents partenaires.

Les points faibles

Du fait de la pseudonymisation des données et donc de l'absence d'accès au NIR (Numéro de sécurité sociale) par Santé publique France et des défauts de qualité d'harmonisation dans la pseudonymisation des bases, il n'a pas été possible d'apparier dès janvier 2021 la base VAC-SI avec les données historiques issues de SI-DEP (base des cas confirmés de Covid-19). Des appariements entre la base VAC-SI et les bases SI-DEP et SI-VIC ont cependant pu être effectués à compter de juillet 2021 à partir d'un pseudonyme commun pour la réalisation notamment d'études d'efficacité vaccinale⁵.

Les données manquantes précédemment décrites, incluant notamment l'absence d'informations initiales sur les départements de résidence des personnes vaccinées, sur leurs activités professionnelles (notamment pour les professionnels de santé) ou le statut de résidents en Ehpad ou plus largement en ESMS et l'absence de visibilité sur la survenue de décès, ont rendu complexe la production d'estimations de couvertures vaccinales de qualité, en particulier dans les populations à risque pour la Covid-19. Par conséquent, des modifications du calcul de l'estimation des CV ont été mises en place, notamment par la prise en compte des nouvelles informations transmises par la Cnam (département de résidence, professionnels de santé...). Du fait de l'imprécision de certaines couvertures vaccinales notamment chez les professionnels de santé et les résidents des établissements médico-sociaux, des études complémentaires ont dû être mises en place^{9,10}.

De plus, l'exploitation quotidienne de l'entièreté de la base de données VAC-SI a également été un véritable challenge en raison de la volumétrie importante et croissante de la base.

Le système VAC-SI, ses évolutions constantes et ses contraintes techniques, ont été très consommateurs et demandeurs de ressources informatiques et en ressources humaines, qui n'étaient pas faciles à obtenir dans des délais très courts entraînant parfois des évolutions tardives. Les réunions de travail avec l'ensemble des partenaires intervenant dans la constitution et la gestion de la base ont néanmoins permis d'échanger et de trouver des solutions rapides à ces problématiques. Ces réunions ont permis de maîtriser l'uniformité des chiffres de couvertures vaccinales communiqués par les différents partenaires.

Vaccin Covid et la base VAC-SI constituent un système spécifique dit « de crise », donc non pérenne. Ce système qui permet le pilotage des campagnes de vaccination contre la Covid-19 a pris fin le 28 juin 2024, sans alternative immédiate. En effet, Vaccin Covid est un outil construit pour la gestion de crise, non interopérable avec les autres outils des professionnels de santé et pour de la vaccination intensive. La maintenance et l'évolution du système représentent des budgets importants. Un tel système pourrait être adapté dans le cadre d'une nouvelle maladie émergente. En revanche, ce système n'est pas adapté au suivi des campagnes de vaccinations liées à la maladie, hors période de crise, puisque les vaccinations contre la maladie Covid-19 ne sont pas remboursées (car fournies gratuitement par l'État français). Des solutions alternatives pour estimer les CV sont en cours d'étude.

Conclusion

La base VAC-SI, alimentée par les informations saisies par les professionnels vaccinateurs dans l'outil Vaccin Covid, développé par la Cnam, a permis de suivre en temps réel dès janvier 2021, et pendant plus de trois ans, l'ensemble des vaccinations contre la Covid-19 réalisées en France par le biais des indicateurs de surveillance, les couvertures vaccinales produites par Santé publique France.

Cette base a été mise en place dans un environnement attestant la sécurité des données des patients et avec des processus de routine automatisés, robustes et réactifs. Le système de surveillance VAC-SI garantit un suivi de qualité des données de vaccinations et dans un contexte de sortie de crise, l'avenir de ce système repose sur la mise en place d'une solution alternative pérenne permettant le suivi des futures campagnes de vaccination et de prise en compte notamment des vaccinations non remboursées car fournies gratuitement, comme ce fut le cas lors de la crise liée à la Covid-19. ■

Remerciements

Les équipes de Santé publique France remercient tous les partenaires ayant participé à l'élaboration et l'évolution du système de surveillance VAC-SI : la Caisse nationale de l'assurance

maladie (Cnam), la Direction du numérique (DNUM), la Direction générale de la santé (DGS), la Direction de la recherche des études de l'évaluation et des statistiques (Drees), les agences régionales de santé (ARS), l'Institut national de la statistique et des études économiques (Insee), et tous les professionnels de santé impliqués dans la surveillance, l'administration des vaccins et la saisie des données.

Liens d'intérêt

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

- [1] Bosetti P, Tran Kiem C, Andronico A, Paireau J, Levy-Bruhl D, Alter L, *et al.* Impact of booster vaccination on the control of COVID-19 Delta wave in the context of waning immunity: Application to France in the winter 2021/22. *Euro Surveill.* 2022;27(1):2101125. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2022.27.1.2101125>
- [2] Vaux S, Blondel C, Platon J, Fonteneau L, Pini A, Hassan E, *et al.* Couverture vaccinale contre la Covid-19 et impact sur la dynamique de l'épidémie. *Bull Épidémiol Hebd.* 2021; (Cov_12):2-13. https://beh.santepubliquefrance.fr/beh/2021/Cov_12/2021_Cov_12_1.html
- [3] Gault G, Bernadou A, Montaufray MA, Filleul L. Impact de la vaccination sur la dynamique de l'épidémie de la Covid-19 chez les personnes âgées de 75 ans et plus en Nouvelle-Aquitaine, janvier-mars 2021. *Bull Épidémiol Hebd.* 2021;(Cov_11):2-8. https://beh.santepubliquefrance.fr/beh/2021/Cov_11/2021_Cov_11_1.html
- [4] Tamandjou Tchuem CR, Auvigne V, Vaux S, Montagnat C, Paireau J, Monnier Besnard S, *et al.* Vaccine effectiveness and duration of protection of COVID-19 mRNA vaccines against Delta and Omicron BA.1 symptomatic and severe COVID-19 outcomes in adults aged 50 years and over in France. *Vaccine.* 2023;41(13):2280-8.
- [5] Tamandjou C, Auvigne V, Schaeffer J, Vaux S, Parent du Châtelet I. Effectiveness of second booster compared to first booster and protection conferred by previous SARS-CoV-2

infection against symptomatic Omicron BA.2 and BA.4/5 in France. *Vaccine.* 2023;41(17):2754-60.

[6] Haute Autorité de santé. Vaccination dans le cadre de la Covid-19. Avis et recommandations. 2024. https://www.has-sante.fr/jcms/p_3178533/fr/vaccination-dans-le-cadre-de-la-covid-19

[7] Figoni J, Campèse C, Spaccaferri G, Rolland P, Caserio-Schönemann C, Che D. Structuration évolutive d'une surveillance multi-sources pour répondre à une infection émergente : l'expérience française face à la Covid-19. *Bull Épidémiol Hebd.* 2023;(1):2-16. https://beh.santepubliquefrance.fr/beh/2023/1/2023_1_1.html

[8] van Kessel R, Forman R, Milstein R, Mastylak A, Czabanowska K, Czypionka T, *et al.* Divergent COVID-19 vaccine policies: Policy mapping of ten European countries. *Vaccine.* 2023;41(17):2804-10.

[9] Santé publique France. Étude de couverture vaccinale contre la COVID-19 chez les professionnels et les résidents des établissements médico-sociaux. Saint-Maurice: Santé publique France; 2021. <https://www.santepubliquefrance.fr/etudes-et-enquetes/etude-de-couverture-vaccinale-contre-la-covid-19-chez-les-professionnels-et-les-residents-des-etablissements-medico-sociaux>

[10] Santé publique France. Étude de couverture vaccinale contre la COVID-19 chez les professionnels et les résidents des établissements médico-sociaux et établissements de santé, saison 2021-2022. Saint-Maurice: Santé publique France; 2022. <https://www.santepubliquefrance.fr/etudes-et-enquetes/etude-de-couverture-vaccinale-contre-la-grippe-et-la-covid-19-dans-les-etablissements-medico-sociaux-et-etablissements-de-sante-saison-2021-2022>

Citer cet article

Platon J, Fonteneau L, Hanguéhard R, Gagnière B, Gault G, Deschamps G, *et al.* VAC-SI : un système d'information pour le suivi de la couverture vaccinale des vaccins contre la Covid-19. *Bull Épidémiol Hebd.* 2024;(20-21):454-61. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_3.html

SURVEILLANCE DE LA PANDÉMIE DE COVID-19 : CONTRIBUTION ET PERFORMANCES DU SYSTÈME SURSAUD®

// SURVEILLANCE OF THE COVID-19 PANDEMIC: CONTRIBUTION AND PERFORMANCE OF THE SURSAUD® SYSTEM

Anne Fouillet¹ (anne.fouillet@santepubliquefrance.fr), Isabelle Pontais¹, Cécile Forgeot¹, Jérôme Naud¹, Gaëlle Pedrono¹, Delphine Serra¹, Bernadette Verrat¹, Groupe des référents SurSaUD® en région*, Patrick Miroux², Pascal Chansard³, Diane Martin⁴, Adel Arfaoui¹, Yann Le Strat¹, Céline Caserio-Schönemann¹

¹ Santé publique France, Saint-Maurice

² Société française de médecine d'urgence (SFMU), Paris

³ SOS Médecins France, Paris

⁴ Centre d'épidémiologie sur les causes médicales de décès de l'Inserm (Inserm-CépiDc), Paris

* Groupe des référents SurSaUD® en région : Nahida Atiki (Cellule régionale Normandie), Marie Barrau (CR Guyane), Oriane Broustal (CR Grand Est), Delphine Casamatta (CR Auvergne-Rhône-Alpes), Sonia Chêne (CR Bourgogne-Franche-Comté), Élise Daudens (CR Antilles), Joël Deniau (CR Provence-Alpes-Côte d'Azur-Corse), Marlène Faisant (CR Bretagne), Noémie Fortin (CR Pays de la Loire), Laure Meurice (CR Nouvelle-Aquitaine), Ali-Mohamed Nassur (CR Océan Indien), Annie-Claude Paty (CR Île-de-France), Leslie Simac (CR Occitanie), Arnoo Shaiykova (CR Hauts-de-France), Marion Soler (CR Mayotte), Nicolas Vincent (CR Centre-Val de Loire).

Soumis le 26.04.2024 // Date of submission: 04.26.2024

Résumé // Abstract

Introduction – Au démarrage de la phase de surveillance populationnelle de la pandémie de Covid-19 en mars 2020, le système de surveillance syndromique SurSaUD® (Surveillance sanitaire des urgences et des décès) constituait un des seuls systèmes déjà en place. Cet article présente une évaluation de l'utilité et des performances du système SurSaUD® pour la surveillance des impacts directs et indirects de l'épidémie de Covid-19 de 2020 à mars 2023.

Méthodes – SurSaUD® collecte quotidiennement, et sans sélection a priori, les données individuelles et pseudonymisées issues des services d'urgences (SU – réseau Oscour®), des associations SOS Médecins (SOSM), des bureaux d'état-civil et des certificats électroniques de décès. L'évaluation a porté sur la flexibilité, l'acceptabilité, la réactivité, la complétude, la stabilité et l'utilité du système au cours de l'épidémie de Covid-19.

Résultats – Les recours aux soins pour suspicion de Covid-19 dans les SU et SOSM, et les décès avec une mention de Covid-19 dans les causes médicales ont évolué de façon concordante avec les indicateurs issus d'autres systèmes de surveillance. Le système a aussi permis de suivre les impacts indirects liés au contexte de la pandémie et de maintenir une surveillance de la mortalité toutes causes et de l'ensemble des recours aux urgences et SOSM. La réactivité et la complétude des données sont restées élevées et stables. Les adaptations du système pour tenir compte de cette émergence (nouveaux codes diagnostics et indicateurs) ont été rapidement mises en place.

Discussion – Afin de renforcer la capacité de surveillance de SurSaUD®, Santé publique France a engagé les démarches pour collecter les données des Samu (Services d'aide médicale urgente). Ce système a été de nouveau confronté à de nombreux défis pour la surveillance des Jeux olympiques et paralympiques de Paris 2024.

Introduction – When the population-based surveillance phase of the COVID-19 pandemic began in March 2020, the syndromic surveillance system SurSaUD® was one of the few systems already in operation. This article assesses the relevance and performance of the SurSaUD® system for monitoring the direct and indirect impacts of the COVID-19 epidemic from 2020 to March 2023.

Methods – SurSaUD® collects pseudonymized individual data on a daily basis and without prior selection from the following sources: the OSCOUR® network of emergency departments (EDs), the SOS Médecins (SOSM) network of general practitioners, civil-status registers and electronic death certificates. The evaluation focused on the system's flexibility, acceptability, responsiveness, completeness, stability and usefulness during the COVID-19 epidemic.

Results – Consultations for suspected COVID-19 in EDs and with SOSMs, as well as death records mentioning COVID-19 as a medical cause, were consistent with the indicators produced by other surveillance systems. The SurSaUD® system also made it possible to monitor the indirect impacts of the pandemic, and to maintain

surveillance of all-cause mortality and all-cause consultations both in EDs and with SOSM. Data responsiveness and completeness remained high and stable. The system was rapidly adapted to integrate new diagnostic codes and indicators.

Discussion – In order to extend the scope of the SurSaUD[®] surveillance system, the French national health agency Santé publique France has begun collecting data from the emergency medical services (SAMU). This system has been confronted to numerous new challenges for monitoring the Paris 2024 Olympic and Paralympic Games.

Mots-clés : Surveillance syndromique, Évaluation, Covid-19, Performance
// **Keywords** : Syndromic surveillance, Assessment, COVID-19, Performance

Introduction

L'année 2020 a été marquée par l'émergence de la Covid-19, qui constitue un épisode sanitaire sans précédent en France comme au niveau mondial depuis le début du ^{xxi}^e siècle, tant sur le plan de son impact sanitaire et sociétal, que sur celui des mesures exceptionnelles de gestion et de prévention qui ont été mises en place pour tenter de ralentir la diffusion du virus dans la population.

Après deux phases de surveillance dédiées à l'identification des premières personnes infectées sur le territoire dès janvier 2020 puis l'investigation des cas groupés, le dispositif a évolué à partir de la mi-mars 2020 vers une surveillance populationnelle, visant à suivre la dynamique spatio-temporelle de l'épidémie et à en évaluer l'impact sur la santé et le système de soins¹. Le système de surveillance sanitaire des urgences et des décès (SurSaUD[®]) constitue un des seuls dispositifs déjà en place pour le démarrage de cette nouvelle phase de la surveillance, avec le réseau Sentinelles, la surveillance des cas graves d'infection respiratoire aiguë hospitalisés dans un réseau de réanimation et la confirmation biologique à travers le système 3-Labos¹. En particulier, le réseau Sentinelles, engagé pour la surveillance de l'épidémie saisonnière grippale au moment de l'émergence de la pandémie, a contribué, dès les premières semaines de l'épidémie, à faire la part entre l'épidémie saisonnière grippale et l'impact de la Covid-19 dans la population, et à identifier les régions les plus touchées par la pandémie².

Développé dès 2004 en réponse à la crise sanitaire de la canicule d'août 2003, le système SurSaUD[®], conçu pour être non spécifique d'une maladie ou d'un risque en particulier, est adaptable et réactif. Construit pour identifier et/ou suivre l'impact sur la population de phénomènes inattendus, connus ou inconnus³, il a déjà été utilisé en situation de pandémie (grippe A(H1N1) en 2009⁴), ainsi que pour le suivi de situations très diverses : émergence infectieuse, épidémies saisonnières hivernales, risques environnementaux ou industriels, attentats, grands rassemblements⁵⁻⁸.

Cet article présente une évaluation des performances et de l'utilité du système SurSaUD[®] pour la surveillance de l'épidémie de Covid-19, de ses impacts directs et indirects, ainsi que des autres indicateurs de routine et saisonniers, de 2020 à mars 2023. Il précise en particulier les modifications qui ont été nécessaires pour adapter le système au contexte de l'émergence.

Matériel et Méthodes

La surveillance syndromique est caractérisée par les principes de collecte et de restitution automatisées et réactives, à partir de données enregistrées dans un système d'information (SI) métier pour d'autres besoins que la surveillance, sans double saisie, donc sans charge de travail supplémentaire pour les partenaires fournisseurs⁹ (dès lors qu'ils disposent d'un logiciel adapté). Le qualificatif « syndromique » vient du type de données recueillies auprès des partenaires fournisseurs : sans aucune sélection a priori car reflétant l'ensemble de l'activité, pas uniquement ou nécessairement médicales, elles peuvent être aussi bien des diagnostics que des signes cliniques ou des symptômes, constituant un diagnostic le plus souvent prévisionnel ou « syndrome ». Toutefois, à la différence de son sens médical strict (tableau clinique ou entité nosologique permettant de porter un diagnostic qui conduit à un traitement), le syndrome peut aussi être, en surveillance, un diagnostic confirmé par des examens complémentaires.

Principes du système de surveillance syndromique SurSaUD[®]

SurSaUD[®] collecte quotidiennement les données individuelles et anonymisées de quatre sources distinctes (tableau 1) :

- les passages dans les services d'urgence (SU) hospitalière participant au réseau Oscour[®] (Organisation de la surveillance coordonnée des urgences) ;
- les actes médicaux (visites à domicile et consultations) des associations SOS Médecins (SOSM) ;
- les décès enregistrés par les bureaux d'état-civil avec transmission automatisée vers l'Institut national de la statistique et des études économiques (Insee) ;
- les décès, dont le volet médical du certificat est transmis par voie électronique au Centre d'épidémiologie sur les causes médicales de décès de l'Institut national de la santé et de la recherche médicale (Inserm-CépiDc).

Les données enregistrées par chaque entité (SU, association SOSM, commune de décès ou médecin certificateur) au cours d'une journée J sont extraites automatiquement des logiciels métiers et transmises chiffrées après minuit le jour suivant (J+1) à Santé publique France, via un concentrateur de données

Tableau 1

Principales caractéristiques des sources de données du système SurSaUD®

Caractéristiques	Oscour®	SOS Médecins	Insee	Inserm-CépiDc ^a
Début de participation au système	2004	2006	2005	2008
Mode de réception des données	1 fichier XML par service d'urgence	1 fichier XML unique pour l'ensemble des associations	1 fichier TXT unique pour l'ensemble des communes	Multiples fichiers XML contenant de 1 à N certificats
Statut de la transmission des données à Santé publique France	2004-2013 : Volontaire Depuis 2014 : transmission obligatoire ^b	Volontaire	Volontaire	Obligatoire ^c
Nombre d'entités participantes : – en mars 2020 – en mars 2023	– 700 services d'urgences – 722 services d'urgences	62 associations (stable entre 2020 et 2023)	– 8 500 communes – 15 062 communes	En 2023 : – environ 1 400 établissements hospitaliers – environ 11 200 médecins
Zones géographiques couvertes	France entière (sauf Martinique)	France métropolitaine et Martinique	France entière	France entière
Couverture en mars 2020 et en mars 2023	93% 95%	95% 95%	90% 95,5%	18% 39%

SurSaUD® : Surveillance sanitaire des urgences et des décès ; Oscour® : Organisation de la surveillance coordonnée des urgences ; Insee : Institut national de la statistique et des études économiques ; Inserm-CépiDc : Centre d'épidémiologie sur les causes médicales de décès de l'Institut national de la santé et de la recherche médicale.

^a certificats électroniques des décès.

^b Voir [10].

^c Voir [11]. L'utilisation de la certification électronique par les médecins est volontaire de 2008 à 2022 et obligatoire pour les décès en établissement hospitaliers et médico-sociaux depuis 2022 [12].

régional (observatoire régional des urgences par exemple) ou national (base SOS Médecins France, Insee ou Inserm-CépiDc) (tableau 1). Un format unique d'extraction des données existe pour l'ensemble des SU (résumé de passage aux urgences – RPU¹⁰), et un autre pour les associations SOSM participant au réseau.

Dans un objectif de consolidation, les fichiers transmis à Santé publique France pour une journée J contiennent respectivement les données des 7 jours (RPU), 3 jours (actes SOSM) et 30 jours (Insee) précédents. Les certificats électroniques de décès peuvent être complétés ou modifiés par le médecin certificateur dans un délai de 96 heures.

Les données des quatre sources comportent des informations démographiques, administratives, et, à l'exception des décès issus de bureaux d'état-civil, des informations médicales :

- le diagnostic principal et le ou les diagnostics associés codés selon la Classification internationale des maladies, 10^e révision (CIM-10) pour les RPU ou selon deux thésaurus spécifiques pour SOSM ;
- les causes médicales inscrites en texte libre par les médecins dans les certificats électroniques de décès.

Les diagnostics Oscour® et SOSM sont analysés sous forme de groupes pertinents définis pour la surveillance : les regroupements syndromiques (RS). L'analyse des causes médicales de décès repose sur des algorithmes basés sur l'analyse du texte libre des certificats.

Indicateurs d'évaluation du système SurSaUD® au cours de l'épidémie de Covid-19

L'évaluation orientée sur les caractéristiques techniques d'un système d'information a porté sur une sélection de six indicateurs, parmi les indicateurs habituels d'évaluation des systèmes de surveillance sanitaire¹³ :

- la flexibilité a été évaluée à travers la capacité du système à intégrer rapidement dans ses référentiels les nouveaux codes diagnostiques dédiés à la Covid-19 (CIM-10 et thésaurus SOSM) et à construire les RS pour suspicion de Covid-19 (« RS Covid-19 ») pour Oscour® et SOSM et un algorithme identifiant les mentions de Covid-19 dans le texte des causes médicales de décès ;
- l'acceptabilité a été mesurée par le nombre de SU ou d'associations SOSM ayant utilisé au moins un des nouveaux codes diagnostiques de Covid-19 dans les premières semaines de l'épidémie. Aux urgences, la dynamique temporelle de la proportion de passages pour chacun des codes composant le RS Covid-19 parmi l'ensemble des passages pour suspicion de Covid-19 (qui seront appelés Covid-19 dans la suite de l'article) est présentée ;
- la complétude des informations a été mesurée d'une part, à partir de la couverture de chacun des réseaux de surveillance, et d'autre part, par la proportion de passages SU (ou actes SOSM) avec un code diagnostique renseigné par rapport à l'ensemble des passages (ou actes reçus) ;

- la réactivité a été mesurée à partir du nombre de passages SU/actes SOSM toutes causes collectés dès J+1 par rapport à ceux collectés au bout de sept jours (données consolidées) ;
- la stabilité de la transmission des données SU et SOSM a été mesurée par le nombre de jours au cours duquel un problème de transmission des données à J+1 a été identifié sur la période de mars à décembre 2020. Ce nombre a été comparé à celui observé sur la période de septembre 2019 à février 2020, précédant le début de l'épidémie ;
- l'utilité du système pour la surveillance des effets directs et indirects de l'épidémie sur la santé de la population est illustrée par : la description de la dynamique des nombres de recours Oscore® et SOSM pour Covid-19 et des décès électroniques avec mention de Covid-19, comparativement au nombre d'hospitalisations pour/avec Covid-19 issu de SI-VIC (Système d'information pour le suivi des victimes d'attentats et de situations sanitaires exceptionnelles, utilisé pour la surveillance intra-hospitalière) ; et l'évolution des indicateurs portant sur les autres pathologies (hors Covid-19) et leur restitution au cours de l'épidémie.

Cette étude ne vise pas une évaluation des performances épidémiologiques des indicateurs de surveillance de la Covid-19 (sensibilité, spécificité, valeur prédictive positive).

Résultats

Flexibilité et acceptabilité du système

Le premier code diagnostique de Covid-19 (U07.1) a été introduit le 24 février 2020 dans les référentiels du système SurSaUD®, soit près de trois semaines après la première recommandation de l'Agence technique de l'information hospitalière (ATIH). Les nouveaux codes, recommandés ensuite par l'ATIH (U07.10 à U07.15) et par SOSM (codes « 278 » et « 218 ») en mars puis en avril, ont été ajoutés dans un délai de deux jours et 24 jours respectivement (tableau 2). Dans le même temps et pour accompagner leur introduction, des recommandations de codage ont été établies conjointement entre Santé publique France, la Fédération des observatoires régionaux des urgences (Fedoru) et la Société française de médecine d'urgences (SFMU), puis diffusées dès le 29 février dans chaque SU, afin d'assurer une utilisation homogène par les professionnels de santé dans leurs logiciels respectifs. La même démarche a été conduite avec la fédération SOS Médecins France.

Tableau 2

Codes des diagnostics médicaux (SU, SOS Médecins) et des motifs (SOS Médecins) en lien avec la Covid-19 ajoutés dans les référentiels utilisés par le système SurSaUD®

	Codes	Libellé	Date de recommandation par l'ATIH	Date d'introduction dans SurSaUD®
Services d'urgences (CIM-10)	B34.2 ^a	Infection coronavirale, sans précision		Avant 2020
	B97.2 ^a	Coronavirus, cause de maladies classées dans d'autres chapitres		Avant 2020
	U04.9 ^a	Syndrome respiratoire aigu sévère (SRAS), sans précision		Avant 2020
	U07.1 ^a	Maladie respiratoire à Coronavirus 2019 (Covid-19)	30/01/2020	24/02/2020
	U07.10 ^a	Covid-19, confirmé	17/03/2020	19/03/2020
	U07.11 ^a	Covid-19, non confirmé	17/03/2020	19/03/2020
	U07.12 ^a	Porteur de SARS-CoV-2 asymptomatique ou pauci symptomatique	17/03/2020	19/03/2020
	U07.13	Autres examens et mises en observations en lien avec l'épidémie Covid-19	17/03/2020	19/03/2020
	U07.14 ^a	Covid-19, autres formes cliniques, virus identifié	10/04/2020	04/05/2020
	U07.15 ^a	Covid-19, autres formes cliniques, virus non identifié	10/04/2020	04/05/2020
SOS Médecins (thésaurus spécifique)	278 (Diagnostic) ^a	Coronavirus	–	Mars 2020
	218 (Motif)	Suspicion de coronavirus	–	Mars 2020
	279 (Diagnostic) ^a	Covid-19 confirmé par la biologie	–	Juillet 2020
	280 (Diagnostic)	Covid-19 personne contact (autres examens et mises en observations en lien avec)	–	Juillet 2020
	281 (Diagnostic)	Nécessité d'une vaccination contre la Covid-19	–	Juillet 2020
	219 (Motif)	Nécessité d'une vaccination contre la Covid-19	–	Juillet 2020

SU : services d'urgences ; SurSaUD® : Surveillance sanitaire des urgences et des décès ; ATIH : Agence technique de l'information sur l'hospitalisation.

^a Codes inclus dans les regroupements syndromiques « suspicion de Covid-19 » pour la surveillance des passages aux urgences et des actes SOSM.

Ces nouveaux codes ont été utilisés par les urgentistes dès leur inclusion dans les logiciels métiers. Le code U07.1 (maladie respiratoire à coronavirus 2019) a été le plus fréquemment utilisé, suivi du code U07.11 (Covid-19, non confirmé)¹⁴ (figure 1). En S10, première semaine de mars 2020, 95 des 700 SU du réseau Oscour® (13,6%) avaient utilisé au moins une fois un des nouveaux codes diagnostiques. Ce nombre a atteint 332 SU (47,4%) en S12 (16/03 au 22/03/2020) et 468 SU (66,8%) en S14 (29 mars au 5 avril 2020). Du côté de SOSM, 32 des 62 SOSM (51,6%) avaient également utilisé au moins une fois le nouveau code diagnostique « 278 » en S10, puis 58 (93,5%) en S12 et 60 (96,8%) en S14. Dès le confinement (S11), le nombre de recours pour suspicion de Covid-19 à partir de ces nouveaux codes était plus fréquent dans les régions Île-de-France, Provence-Alpes-Côte d'Azur et Grand Est.

En parallèle, deux nouveaux RS Covid-19 ont été créés et leur définition a été élargie au fur et à mesure de l'introduction des nouveaux codes (tableau 2). Les indicateurs correspondant aux RS Covid-19 Oscour® et SOSM ont été présentés dans le point épidémiologique (PE) Covid-19 à partir du 15 mars 2020⁽¹⁾ au niveau national et du 10 avril dans les PE régionaux, et dans les bulletins hebdomadaires de routine dédiés à l'activité des sources Oscour® et SOSM dès le 17 mars¹⁵. Ils ont également été publiés

à disposition du grand public et mis à jour quotidiennement sur l'observatoire cartographique Géodes à partir du 20 mars 2020¹⁶.

En ce qui concerne les causes de décès, un algorithme d'identification des causes médicales contenant une mention de Covid-19 a été développé de façon conjointe entre Santé publique France et le CépiDc. Afin d'évaluer les éventuelles comorbidités des personnes décédées de (ou avec) la Covid-19, des algorithmes ont également été développés pour identifier les mentions d'obésité, diabète, pathologies cardiaques, rénales, respiratoires ou neurodégénératives/neurocognitives. Ces indicateurs ont été introduits dans le PE Covid-19 à partir du 24 mars 2020.

Complétude, réactivité, stabilité

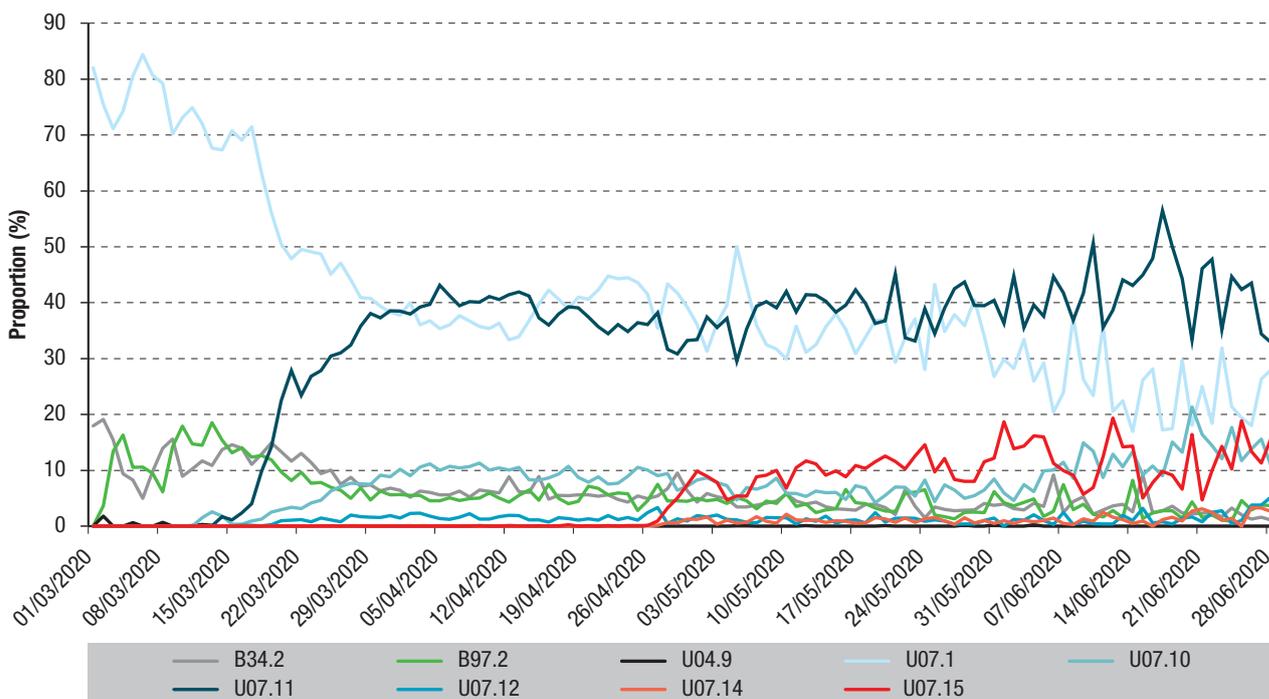
La couverture des réseaux Oscour® et SOSM, supérieure à 93% en mars 2020, est restée stable au cours de l'épidémie (tableau 1). En ce qui concerne la mortalité, le nombre de communes transmettant automatiquement les données de mortalité à l'Insee a connu une forte progression au début de la pandémie, passant d'environ 8 500 communes en mars 2020 (90% de la mortalité nationale) à près de 10 500 communes en mai 2020 (93% de la mortalité). Fin mars 2023, ce nombre atteignait près de 15 100 communes (95%). Parallèlement, 18% des décès étaient certifiés par voie électronique au début de la pandémie. Ce taux de déploiement a rapidement progressé jusqu'en mai 2020, atteignant 26%, puis de façon beaucoup plus lente, pour atteindre 39% en mars 2023.

Quant à la réactivité, 93,0% des passages aux urgences toutes causes ont été reçus et intégrés

(1) Santé publique France. Point Épidémiologique Covid-19. <https://www.santepubliquefrance.fr/recherche/#search=COVID%2019%20%20%20point%20epidemiologique&publications=donn%C3%A9es®ions=National&sort=date>

Figure 1

Proportion quotidienne du nombre de passages aux urgences pour chacun des diagnostics inclus dans le regroupement « Suspicion de Covid-19 » parmi l'ensemble des passages pour suspicion de Covid-19, France, du 1^{er} mars au 28 juin 2020



La signification des codes est disponible dans le tableau 2.

dès le lendemain du passage (J+1) entre mars et décembre 2020. Parmi eux, 72,5% contenaient un diagnostic principal codé (figure 2), passant à 79,2% dès J+3 avec la consolidation des données. Ces proportions ont légèrement progressé sur cette période, comparativement aux six mois précédant la pandémie (91,0%, 70,5% et 77,4% respectivement).

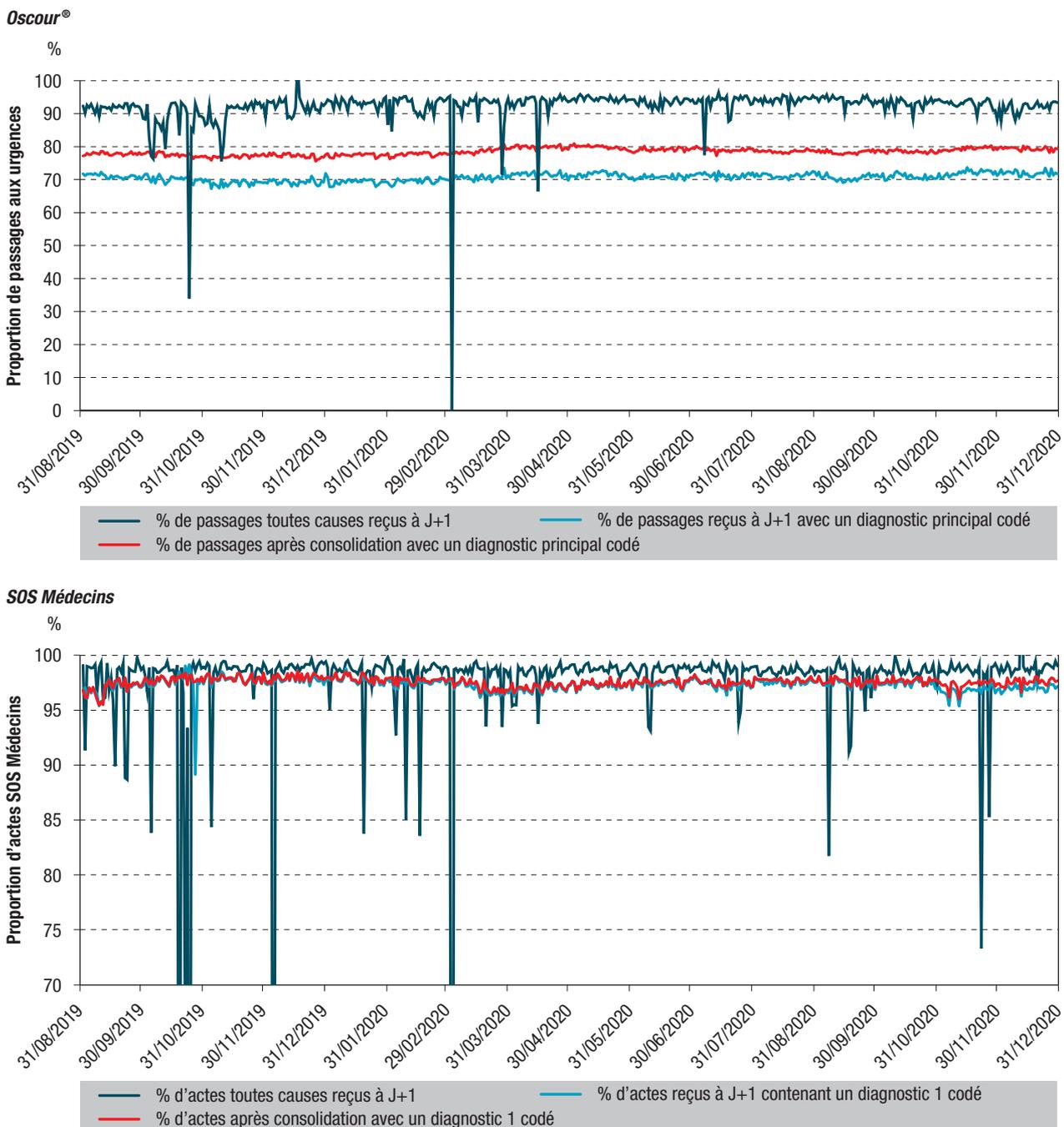
Sur la même période, parmi les actes SOSM toutes causes, 98,0% ont été reçus et intégrés à J+1, en légère hausse par rapport aux six mois précédents (96,4%). Parmi les actes reçus à J+1, 97,2% contenaient au moins un diagnostic médical codé (figure 2).

Cette proportion d'actes avec un diagnostic codé atteignait 97,7% dès J+2.

La réactivité est restée stable, quels que soient le jour de la semaine et la période de l'année (figure 2). Si, ponctuellement sur la période d'étude, la proportion de passages aux urgences reçus à J+1 a diminué de façon marquée, cela a concerné un nombre restreint de journées : 1 journée (le 3 mars 2020) pour problème technique, 4 journées entre mars et novembre 2020 pour absence de transmission de données d'une ou plusieurs régions. Ces difficultés ont toutes été résolues dans la journée.

Figure 2

Proportions quotidiennes du nombre de passages aux urgences et d'actes SOS Médecins toutes causes, France, septembre 2019 à décembre 2020



Du côté de SOSM, si le nombre de difficultés ponctuelles s'est fortement réduit entre mars et décembre 2020, l'agence a dû faire face à une interruption de transmission de 22 associations SOSM (30% des actes) entre mi-février et le 6 juillet 2022, à la suite d'une évolution technique sur les serveurs de SOSM. Sur cette période, la surveillance sanitaire a été maintenue sur les 39 autres associations. Les données de ces associations ont pu être récupérées a posteriori.

En ce qui concerne les deux sources de mortalité, les transmissions des données sont également restées stables tout au long de la période de la pandémie (résultats non présentés).

Utilité du système pour la surveillance de la Covid-19

Avant même la phase de surveillance populationnelle et la création des codes Covid-19 par l'ATIH, plusieurs RS Oscour® et SOSM, traceurs des symptômes cliniques de la Covid-19 (fièvre isolée, toux, dyspnée/insuffisance respiratoire et grippe/syndrome grippal) étaient déjà disponibles et suivis en routine dans le cadre de la surveillance de l'épidémie saisonnière grippale. Ils ont alors été également analysés pour tenter d'identifier les premiers impacts de la Covid-19 sur le recours aux soins.

Après une augmentation des recours pour fièvre isolée, corrélée à celle des recours pour grippe/syndrome grippal (entre S50-2019 et S10-2020), une reprise modérée des recours pour fièvre isolée est à nouveau observée à partir de S11-2020 aux urgences et chez SOSM. Cette hausse était contemporaine d'une augmentation des recours pour toux dans les deux réseaux, particulièrement marquée dans les SU (figure 3) ; les recours pour toux sont ainsi passés de 15 737 actes SOSM en S10-2020 à 18 912 actes en S11-2020 (soit +20%), et de 2 730 passages aux urgences en S10-2020 à 6 533 passages en S12-2020 (soit +139%). Une hausse des recours aux urgences pour dyspnée/insuffisance respiratoire était également notée sur la période (+61% entre S10-2020 et S13-2020).

Avec la création des codes spécifiques Covid-19 et la diffusion des recommandations de codage dans les réseaux, les recours pour toux et fièvre isolée ont rapidement diminué, ainsi que, dans une moindre mesure, ceux pour dyspnée/insuffisance respiratoire (figure 3). Entre les semaines S10-2020 et S13-2023, 544 834 actes SOSM et 936 437 passages aux urgences pour le RS Covid-19 ont été enregistrés, représentant respectivement 4,9% et 1,9% de l'activité totale codée des SU et de SOSM sur l'ensemble de la période. Au pic de la première vague (S13-2020), ces recours ont atteint 20% de l'activité totale codée des SU et 24% de celle de SOSM. Aux pics des vagues suivantes (S44-2020, entre S12-2021 et S14-2021, S32-2021, entre S01-2022 et S03-2022, S14-2022 et S27-2022), ces proportions étaient comprises entre 2 et 7% de l'activité des urgences et entre 5% et 20% de l'activité SOSM.

La dynamique hebdomadaire du nombre des recours aux urgences pour Covid-19 était concordante avec celle du nombre d'actes SOSM (coefficient de corrélation=0,79). La dynamique de ces deux indicateurs était également concordante avec celle du nombre d'hospitalisations pour ou avec Covid-19 enregistrées à partir du système SI-VIC (coefficient de corrélation=0,89 entre Oscour® et SI-VIC, et 0,74 entre SOSM et SI-VIC) (figure 4).

En ce qui concerne la mortalité, entre S10-2020 et S13-2023, 82 202 certificats électroniques de décès avec une mention de Covid-19 dans les causes médicales ont été enregistrés par le système. La dynamique temporelle était également cohérente, avec un décalage d'une semaine, avec celle des hospitalisations pour ou avec Covid-19 (SI-VIC) et des recours aux urgences et à SOSM pour Covid-19 (figure 4). Ainsi, le coefficient de corrélation entre le nombre de passages aux urgences pour Covid-19 et le nombre de décès de la semaine suivante était de 0,83. Cette corrélation était plus forte en considérant le nombre d'hospitalisations après passage aux urgences pour Covid-19 ou les hospitalisations pour ou avec Covid-19 dans SI-VIC et le nombre de décès de la semaine suivante (0,91 et 0,90 respectivement).

Surveillance des recours aux soins pour les autres pathologies pendant l'épidémie

En parallèle à la surveillance de la Covid-19, le suivi quotidien des indicateurs habituels de surveillance syndromique (morbidité comme mortalité) a été maintenu tout au long de l'épidémie.

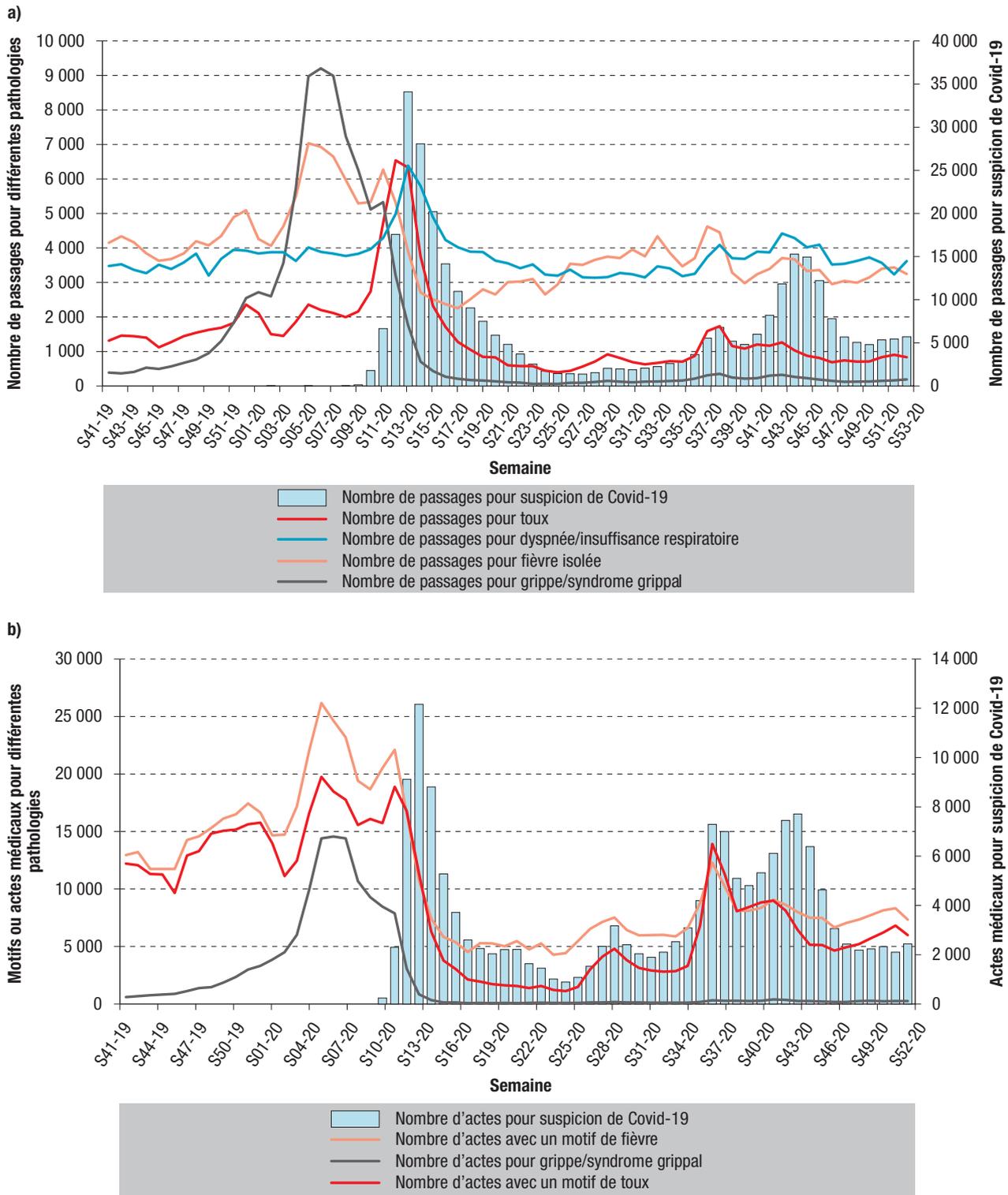
L'activité toutes causes a fortement baissé pendant le premier confinement (du 16 mars au 10 mai 2020 : -45% aux urgences, -32% chez SOSM). En cohérence avec cette tendance, les recours aux urgences pour infarctus du myocarde et accident vasculaire cérébral (AVC) ont également baissé, principalement pendant le premier confinement (respectivement -24% et -18% par rapport aux périodes équivalentes entre 2017 et 2019) ; le suivi de ces recours a permis de montrer une absence de rebond immédiat après le confinement, malgré la crainte d'un retard de prise en charge pour ces pathologies graves¹⁷.

De même, la surveillance de routine a contribué à objectiver l'impact de la pandémie sur la santé mentale de la population. Ainsi, les actes SOSM pour angoisse chez les adultes de 15-74 ans ont progressé au cours du premier confinement, suivis à partir de novembre 2020, par les passages aux urgences pour idées suicidaires et troubles de l'humeur chez les enfants. Ces observations précoces, confortées par des signalements de professionnels fin 2020 puis début 2021, ont conduit à mettre en place un PE dédié à la surveillance de la santé mentale à partir du 2 février 2021. Ce PE a été produit de façon hebdomadaire jusqu'au 18 juillet 2022, avant de passer à un rythme mensuel⁽²⁾. Une déclinaison régionale

⁽²⁾ Santé publique France. Bulletin Surveillance Syndromique Santé Mentale. <https://www.santepubliquefrance.fr/recherche/#search=Bulletin%20surveillance%20syndromique%20sant%C3%A9%20mentale>

Figure 3

Nombre hebdomadaire de recours aux urgences (a), motifs d'appels et actes médicaux SOS Médecins (b) pour des symptômes évocateurs de la Covid-19, France, octobre 2019 à décembre 2020



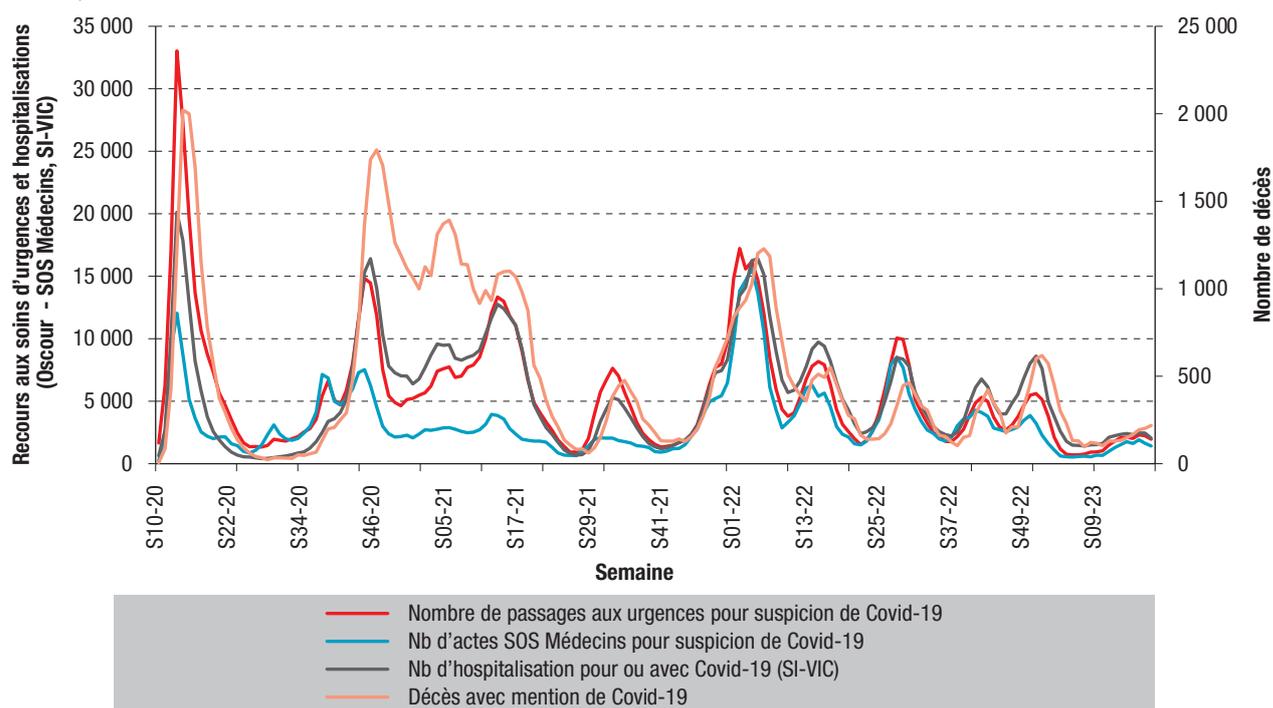
de ce PE a également été diffusée trimestrielle- ment. Ces indicateurs montraient toujours en 2023 des niveaux élevés persistants des recours pour idées suicidaires principalement chez les jeunes.

Les surveillances saisonnières (hivernale : infections respiratoires aiguës ; à prédominance printanière : méningite virale, allergie ; estivale : pathologies liées

à la chaleur, noyade, asthme de la rentrée scolaire), qui reposent pour tout ou partie sur les données Oscour® et SOSM, ont également été assurées tout au long de la période pandémique. Ainsi, la forte intensité de l'épidémie saisonnière de bronchiolite en période post-Covid (saison 2022-2023) a pu être objectivée et faire l'objet d'une surveillance renforcée, centrée sur les moins de 1 an, plus fragiles. Les bulletins

Figure 4

Nombre hebdomadaire de recours aux urgences, d'actes médicaux SOS Médecins pour suspicion de Covid-19 et de décès électroniques avec une mention de Covid-19, et nombre d'hospitalisations avec ou pour Covid-19 (SI-VIC), France, mars 2020 à mars 2023



hebdomadaires pour l'ensemble des indicateurs Oscour® et SOSM ont été publiés sans modification de méthode ou de fréquence d'analyse.

L'analyse de la mortalité toutes causes confondues à partir des données Insee a également été poursuivie, permettant de prendre en compte, sans pouvoir les distinguer, les effets directs de la Covid-19, ceux liés au contexte de la pandémie (notamment réduction des déplacements), ainsi que les effets d'autres événements intercurrents (vagues de chaleur, épidémies saisonnières hivernales). L'estimation de l'excès de mortalité lors de la première vague a ainsi pu être produite rapidement, en mesurant l'écart entre le nombre observé et le nombre attendu de décès, estimé à partir du modèle statistique utilisé par le consortium EuroMomo¹⁸. Le bulletin hebdomadaire national de surveillance de la mortalité toutes causes a été temporairement interrompu entre le 24 mars 2020 et le 2 février 2021, les informations étant directement restituées chaque semaine dans le PE Covid-19.

Les décès certifiés par voie électronique, pour lesquels les causes médicales ne contenaient pas de mention de Covid-19, ont contribué à interpréter les évolutions de la mortalité toutes causes et à suivre la mortalité pour d'autres causes que la Covid-19¹⁹. En particulier, cette source a permis de faire une première évaluation de la mortalité par suicide au cours de la pandémie²⁰.

Discussion

Le système multisource SurSaUD®, conçu et adapté pour assurer la surveillance sanitaire de tout type d'événement sanitaire, a constitué l'un des rares

outils immédiatement disponibles et opérationnels aux premières heures de la pandémie de Covid-19.

Les performances du dispositif sont restées remarquables tout au long de la crise. Malgré l'impact massif de la pandémie sur le système de soins, notamment hospitalier, les taux de transmission des données des urgences et des associations SOSM vers le système SurSaUD® (93,0% des passages aux urgences et 98,0% des actes SOSM reçus dès le lendemain à Santé publique France) et les taux de codage des diagnostics médicaux dans un délai de 24 heures (72,5% aux urgences et 97,2% des actes SOSM) sont restés stables tout au long de l'année 2020 et ont même légèrement progressé par rapport aux mois précédents. Il a également pu être mis en place rapidement la transmission de données pour certaines unités spécifiques Covid-19 créées « à côté » de certains établissements hospitaliers lors du premier confinement, lorsqu'ils utilisaient le format RPU. Bien qu'il puisse y avoir une hétérogénéité d'un établissement à l'autre, les nouveaux codes diagnostiques ont été utilisés par les professionnels de santé dans les trois jours qui ont suivis la recommandation de l'ATIH, témoignant d'une grande adaptabilité de la plupart des logiciels métiers et de l'adhésion des médecins aux recommandations de codage ; ils ont pu être rapidement intégrés dans les référentiels de l'agence pour permettre la construction de nouveaux RS dans un délai court (deux jours maximum après leur intégration). Le délai plus long de 24 jours qui a pu être observé en mai 2020 était lié 1/ au délai pour identifier ou recevoir les recommandations de l'ATIH, 2/ au besoin d'échanges avec les médecins des réseaux sur leur utilisation de ces codes et

3/ à l'évaluation de l'impact de l'ajout des nouveaux codes dans les RS sur la continuité des indicateurs et à l'accompagnement de ces changements auprès des utilisateurs (épidémiologistes, décideurs, médias).

Les recours aux soins pour Covid-19 dans les SU et les associations SOSM, comme les décès avec une mention de Covid-19 dans les causes médicales, ont évolué au niveau national de façon concordante avec les indicateurs produits par les autres systèmes de surveillance spécifiques (coefficients de corrélation compris entre 0,74 et 0,90 selon les sources), constituant un argument fort de validité du dispositif pour le suivi de la dynamique d'un phénomène épidémique émergent de grande ampleur comme la Covid-19. Lors du premier confinement, les indicateurs Oscour® et SOSM ont par ailleurs été les premiers, parmi les indicateurs existants, à montrer le début de décroissance de la première vague à partir du 3 avril 2020.

Au-delà des indicateurs directement liés à la Covid-19, le système a permis :

- de suivre les impacts indirects de la pandémie tels que la réduction des recours aux soins pendant les confinements. En particulier, la baisse du recours aux urgences pour infarctus du myocarde et pour AVC objectivé avec cet outil¹⁷ a conduit à des recommandations à la population et des études complémentaires, notamment sur les hospitalisations et la mortalité cardiovasculaire pendant le début de la pandémie²¹ ;
- de mettre en évidence une augmentation des recours pour des indicateurs de santé mentale, en particulier chez les enfants. Pour éclairer la décision publique, Santé publique France a également mis en œuvre des études et enquêtes spécifiques pour renforcer la surveillance de la santé mentale en population générale (CoviPrev²²) et dans des populations spécifiques (Enabee²³, Confeado²⁴, Covimater²⁵, étude CovSA, cohorte Coset...);
- de maintenir la surveillance de routine sur tout le reste du champ, en morbidité comme en mortalité, en particulier les épidémies saisonnières (bronchiolite, gastro-entérite, grippe, etc.) ou les effets des événements climatiques (vagues de chaleur des étés 2020 à 2023).

Forces du système

Le caractère multisource du système SurSaUD®, couvrant médecine d'urgence de ville et hospitalière ainsi que la mortalité, assure une analyse croisée et complémentaire des situations sanitaires en essayant de réduire les angles morts, sans toutefois les supprimer totalement. Il tend également à rendre la surveillance moins sensible à des décisions politiques ou organisationnelles, pouvant entraîner une modification des comportements de recours aux soins ou de prise en charge de la population, ou des difficultés de remontée de données (grève de transmission, problème ou évolution technique...).

L'une des principales forces du système réside dans l'utilisation de données déjà collectées en routine par les réseaux de professionnels pour leurs propres besoins, sous réserve de disposer de logiciels métiers adaptés. Avec l'automatisation de l'extraction des données des logiciels et de la transmission vers Santé publique France, le système n'impose ainsi aucune charge de travail supplémentaire aux professionnels de santé pour contribuer à la surveillance, à la différence de systèmes non intégrés comme SI-VIC qui ont occasionné une surcharge de double saisie pour les établissements déjà très impactés. Cette intégration est un facteur favorisant la stabilité de la transmission. Elle reste toutefois conditionnée à la capacité de saisie des RPU par les professionnels de santé, en particulier dans des situations où l'organisation des soins est bouleversée. Par ailleurs, les aménagements nécessaires à Santé publique France pour adapter le système au contexte de l'émergence (ajout des nouveaux codes diagnostics dans les référentiels et la création des nouveaux regroupements syndromiques dédiés à la Covid-19) ont été très limités, correspondant à des fonctionnalités déjà opérationnelles dans le système d'information.

Outre la technique, l'autre grande force du système réside dans les réseaux humains Oscour® et SOSM, impliqués et réactifs, disposant à la fois d'une bonne connaissance du terrain et du système de surveillance. Dès le début de la pandémie, les premiers échanges ont permis d'établir conjointement, puis d'assurer la diffusion rapide des recommandations de codage auprès de l'ensemble des SU et associations SOSM. Des réunions hebdomadaires puis bimensuelles ont été organisées, afin de confronter les observations épidémiologiques et l'analyse clinique des professionnels de santé sur le terrain et de présenter aux partenaires l'évolution des indicateurs. Ces échanges ont aidé à une meilleure compréhension des situations locales, à l'interprétation des indicateurs pour Santé publique France et à donner un espace de discussion aux professionnels pour partager leurs ressentis et leurs expériences.

L'antériorité du système, qui collecte des données depuis 2004, a permis à la fois de disposer d'une large couverture (à l'exception de la certification électronique des décès) sur l'ensemble du territoire, incluant les territoires ultramarins et d'une bonne connaissance préalable des données, des indicateurs et de leur interprétation. Les données Oscour® ont été utilisées pour mesurer et suivre le taux de reproduction (R-effectif) de l'épidémie de Covid-19, marqueur de la diffusion du virus dans la population, en complément des R-effectifs calculés à partir des données de tests (SI-DEP) et d'hospitalisations (SI-VIC)²⁶.

Par ailleurs, grâce à la couverture territoriale du réseau, les indicateurs Oscour® ont contribué directement à la prise de décision, pour préparer la levée progressive du premier confinement. Ainsi, en l'absence de système de surveillance virologique basé sur les tests de dépistage (SI-DEP en cours

de construction²⁷), l'un des trois indicateurs utilisés pour construire la carte départementale de déconfinement jusqu'au 28 mai 2020 reposait sur les recours aux urgences pour Covid-19, qui permettaient de définir le niveau de circulation active du virus²⁸ au plus proche de la contamination.

Enfin, la diversité des variables remontées pour chacune des sources a permis d'affiner les analyses pour certaines classes d'âges et certains diagnostics associés. Ainsi, en complément des données de vaccination, les indicateurs Oscour[®] et SOSM ont été utilisés à la rentrée scolaire 2021²⁹ pour dresser un état des lieux des contaminations et suivre l'évolution de la Covid-19 chez les enfants et les étudiants avec le retour en collectivité. Début 2022, les recours aux urgences ont également contribué à décrire et caractériser l'évolution des symptômes et diagnostics associés au variant Omicron, en mettant en évidence une diminution des formes respiratoires, prédominantes pendant la circulation du variant Delta notamment³⁰, au profit d'une augmentation des formes non respiratoires de la Covid-19.

Limites du système

Le déploiement limité de la certification électronique des décès au démarrage de l'épidémie (18% de la mortalité nationale, principalement utilisée par les établissements hospitaliers) a constitué la principale limite du système, ne permettant pas la mise en œuvre systématique d'une surveillance réactive de la mortalité par cause et la caractérisation des comorbidités des décès en lien avec la Covid-19. Sa faible montée en charge au cours de l'épidémie (35% de la mortalité nationale mi-2022), malgré des instructions ministérielles auprès des établissements hospitaliers et médico-sociaux pour prioriser l'utilisation du système, pourrait être le reflet de la concurrence des systèmes spécifiques mis en place pour la surveillance de l'épidémie de Covid-19 et intégrant des indicateurs de mortalité (resp. SI-VIC et SurvESMS).

Si les données issues des SU et des associations SOSM permettent une surveillance de la dynamique d'un événement sanitaire et la description des patients (classe d'âge, sexe, sévérité), elles ne permettent pas de disposer d'informations sur l'état de santé des patients (antécédents, comorbidités ou statut vaccinal) au moment du recours. De plus, l'absence d'identifiant unique de santé ne permet pas non plus d'envisager le chaînage de ces données avec d'autres sources remontées de façon réactive et pouvant venir les compléter, telles que VAC-SI (suivi de la vaccination)³¹ ou SI-DEP (résultats des tests biologiques)²⁷.

Enfin, pour des raisons de limitation technique des outils, l'analyse des recours est réalisée sur la base de la zone géographique de consultation, ce qui ne permet pas de produire des taux d'incidence. L'utilisation de la commune de résidence des patients, présente dans les données transmises par les SU comme SOSM, constitue une perspective

d'amélioration de la surveillance, notamment pour les événements impliquant des flux de populations (vacances, grands rassemblements de population,...) ou des populations exposées en cas d'accidents industriels par exemple.

Éléments de comparaison internationale

Au niveau international, le système de surveillance syndromique du Royaume-Uni constitue le système multisource le plus complet et avancé. Il combine une surveillance de la médecine de ville, des recours aux SU (basés sur le modèle Oscour[®]), des appels vers une plateforme de conseil et d'orientation médicale, ainsi que vers les plateformes de régulation d'ambulances³². À l'instar de SurSaUD[®], le système a apporté une contribution majeure pour le suivi de l'impact direct et indirect de la pandémie sur les recours aux soins³³. Il a ainsi mis en évidence une diminution majeure des passages dans les services d'urgences entre les semaines S13-2020 et S20-2020, comme en France, et à l'inverse, l'accroissement marqué des appels aux plateformes de conseil et d'orientation médicale et de régulation d'ambulances, reflétant ainsi le changement de pratique de la population face à l'inquiétude de la contamination en cas de consultation.

De nombreux autres pays disposent d'une surveillance des recours aux urgences toutes causes et pour différentes pathologies hors Covid-19 et ont pu mettre en évidence dans les premiers mois de l'épidémie des tendances similaires à celles observées en France^{33,34}.

Aux États-Unis, la pandémie de Covid-19 a permis d'améliorer, d'une part, le déploiement du système de surveillance syndromique fondé sur les SU, atteignant en 2023 une couverture de 75% des recours nationaux et, d'autre part, la réactivité de la remontée des certificats de décès (incluant les causes médicales) au système statistique des statuts vitaux³⁵. Une évaluation comparative des différents systèmes utilisés pour la surveillance de la Covid-19 a montré que les recours aux urgences pour Covid-19 ont permis d'identifier les tendances avec une avance de quatre jours par rapport aux admissions hospitalières³⁶. De même, les certificats de décès assuraient une surveillance des tendances de la mortalité plus réactive que le système spécifique mis en place pour la Covid-19, reposant sur la collecte de données agrégées.

Perspectives

Afin de renforcer la capacité de surveillance du système SurSaUD[®], Santé publique France a engagé les démarches pour collecter les données des Samu. Le déploiement massif du système de certification électronique des décès doit également être fortement encouragé sans attendre une nouvelle crise sanitaire. Un déploiement en cours de crise sanitaire, qui n'atteint pas ou ne s'approche pas de l'exhaustivité ou de la représentativité, représente plus une difficulté d'analyse qu'un avantage. D'autres chantiers pour renforcer le dispositif de surveillance multisource dans le contexte

pandémique sont également ouverts par Santé publique France³⁷, en particulier le projet Orchidée visant à tester la production réactive d'indicateurs spécifiques à partir des entrepôts de données de santé des établissements hospitaliers.

Au lendemain de la pandémie de Covid-19, le système SurSaUD® a été de nouveau confronté à de nombreux défis pour la surveillance des Jeux olympiques et paralympiques de Paris 2024. L'afflux important de population du monde entier pour cet événement international et festif a pu entraîner des risques sanitaires spécifiques aux grands rassemblements de population (contamination alimentaire, traumatismes liés à des mouvements de foule, malaises, intoxications (alcools, drogues), menaces terroristes...), qui s'ajoutent aux menaces sanitaires habituelles de la période estivale (vague de chaleur, orages violents, incendies de forêt, sécheresse, noyades, extension des arboviroses (chikungunya, dengue, zika)). ■

Remerciements

Les auteurs remercient l'ensemble des urgentistes et des médecins des associations SOS Médecins et certificateurs de décès, ainsi que les personnels de l'Insee (en particulier Sylvie Le Minez, Chantal Villette, Valérie Roux, Anne Levet et Agnès Lerenard) et du Cépidec-Inserm (en particulier, Grégoire Rey, Aude Robert, Pierre-Etienne Alary) impliqués dans la constitution des données et de leur transmission vers Santé publique France, ainsi que pour les échanges tout au long de l'épidémie de Covid-19 pour la constitution et l'interprétation des indicateurs. Les auteurs remercient également l'ensemble des partenaires impliqués de la Fedoru et de la SFMU.

Les auteurs remercient également Jérôme Guillevic pour sa relecture attentive de l'article et Minh-Canh Quan pour son appui majeur pour le maintien en condition opérationnelle de l'application informatique associée au système SurSaUD®, ainsi que les référents SurSaUD® en région : Audrey Andrieu, Valentin Courtillet, Erica Fougère, Céline François, Quiterie Mano et Jérôme Pouey.

Liens d'intérêt

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

- [1] Figoni J, Campèse C, Spaccaferri G, Rolland P, Caserio-Schönemann C, Che D. Structuration évolutive d'une surveillance multi-sources pour répondre à une infection émergente : l'expérience française face à la Covid-19. *Bull Épidémiol Hebd.* 2023;(1):2-16. http://beh.santepubliquefrance.fr/beh/2023/1/2023_1_1.html
- [2] Boëlle PY, Souty C, Launay T, Guerrisi C, Turbelin C, Behillil S, *et al.* Excess cases of influenza-like illnesses synchronous with coronavirus disease (COVID-19) epidemic, France, March 2020. *Euro Surveill.* 2020;25(14):2000426. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.14.2000326>
- [3] Caserio-Schönemann C, Bousquet V, Fouillet A, Henry V, pour l'équipe projet SurSaUD®. Le système de surveillance syndromique SurSaUD®. *Bull Épidémiol Hebd.* 2014;(3-4):38-44. http://beh.santepubliquefrance.fr/beh/2014/3-4/2014_3-4_1.html
- [4] Jossieran L, Caillère N, Goncalves N, Ringard D, Leroy C, Fournet N, *et al.* Surveillance syndromique dans le cadre de la pandémie grippale A(H1N1)2009 : intérêts et limites. *Bull Épidémiol Hebd.* 2010;(24-25-26):274-7. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/documents/article/surveillance-syndromique-dans-le-cadre-de-la-pandemie-grippale-a-h1n1-2009-interets-et-limites>
- [5] Fouillet A, Fournet N, Forgeot C, Jones G, Septfons A, Franconeri L, *et al.* Large concomitant outbreaks of acute gastroenteritis emergency visits in adults and food-borne events suspected to be linked to raw shellfish, France, December 2019 to January 2020. *Euro Surveill.* 2020;25(7):2000060. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.7.2000060>
- [6] Pelat C, Bonmarin I, Ruello M, Fouillet A, Caserio-Schönemann C, Levy-Bruhl D, *et al.* Improving regional influenza surveillance through a combination of automated outbreak detection methods: The 2015/16 season in France. *Euro Surveill.* 2017;22(32):30593. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2017.22.32.30593>
- [7] Vandentorren S, Paty AC, Baffert E, Chansard P, Caserio-Schönemann C. Syndromic surveillance during the Paris terrorist attacks. *Lancet.* 2016;387(10021):846-7.
- [8] Marguerite N, Brottet E, Pages F, Jaffar-Bandjee MC, Schuffenecker I, Josset L, *et al.* A major outbreak of conjunctivitis caused by coxsackievirus A24, Réunion, January to April 2015. *Euro Surveill.* 2016;21(26). <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2016.21.26.30271>
- [9] Triple S Project. Assessment of syndromic surveillance in Europe. *Lancet.* 2011;378(9806):1833-4.
- [10] Ministère des Affaires sociales et de la Santé. Arrêté du 24 juillet 2013 relatif au recueil et au traitement des données d'activité médicale produites par les établissements de santé publics ou privés ayant une activité de médecine d'urgence et à la transmission d'informations issues de ce traitement dans les conditions définies à l'article L. 6113-8 du code de la santé publique et dans un but de veille et de sécurité sanitaires. *JORF.* 2013;(0185):9-11. <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000027825549>
- [11] République Française. Article R2213-1-3 du Code général des collectivités territoriales. https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000034480102/2024-09-18/
- [12] Ministère des Solidarités et de la Santé. Décret n° 2022-284 du 28 février 2022 relatif à l'établissement du certificat de décès. *JORF.* 2022;(0050):128-9. <https://www.legifrance.gouv.fr/download/pdf?id=yolCDXjzuZbzb6fvzvj77DIKj9JHfFlb5CypBFTWY5w>
- [13] German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN. Updated guidelines for evaluating public health surveillance systems: Recommendations from the Guidelines Working Group. *MMWR Recomm Rep.* 2001;50(RR-13):1-35.
- [14] Thiam MM, Pontais I, Forgeot C, Pedrono G, SurSaUD® Regional Focal Point, SOS Médecins, *et al.* Syndromic surveillance: A key component of population health monitoring during the first wave of the COVID-19 outbreak in France, February-June 2020. *PLoS One.* 2022;17(2):e0260150.
- [15] Santé publique France. Bulletin national d'information OSCOUR du 17 mars 2020. Saint-Maurice: Santé publique France; 2020. 18 p. <https://www.santepubliquefrance.fr/surveillance-syndromique-sursaud-r/documents/bulletin-national/2020/bulletin-national-d-information-oscour-du-17-mars-2020>
- [16] Lucas E, Fouquet A, Jezewski-Serra D, Ben Raies J, De Crouy-Chanel P, Alleaume C. Santé publique France face au défi de l'open data. *Bull Épidémiol Hebd.* 2024;(20-21):475-81. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_5.html
- [17] Olié V, Carcaillon-Bentata L, Thiam MM, Haeghebaert S, Caserio-Schönemann C. Emergency department admissions for myocardial infarction and stroke in France during the first wave of the COVID-19 pandemic: National temporal trends and regional disparities. *Arch Cardiovasc Dis.* 2021;114(5):371-80.

- [18] Fouillet A, Pontais I, Caserio-Schönemann C. Excess all-cause mortality during the first wave of the COVID-19 epidemic in France, March to May 2020. *Euro Surveill.* 2020; 25(34):2001485. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.34.2001485>
- [19] Santé publique France. À quoi peut-on attribuer la hausse de la mortalité observée depuis novembre 2021 ? Point au 10 février 2022. Saint-Maurice: Santé publique France; 2022. 11 p. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/enquetes-etudes/a-quoi-peut-on-attribuer-la-hausse-de-la-mortalite-observee-depuis-novembre-2021-point-au-10-fevrier-2022>
- [20] Fouillet A, Martin D, Pontais I, Caserio-Schönemann C, Rey G. Reactive surveillance of suicides during the COVID-19 pandemic in France, 2020 to March 2022. *Epidemiol Psychiatr Sci.* 2023;32:e20.
- [21] Gabet A, Grave C, Tuppin P, Chatignoux E, Bejot Y, Olié V. Impact of the COVID-19 pandemic and a national lockdown on hospitalizations for stroke and related 30-day mortality in France: A nationwide observational study. *Eur J Neurol.* 2021;28(10):3279-88.
- [22] Santé publique France. Comment évolue la santé mentale des Français ? Résultats de la vague 37 de l'enquête Covi-Prev (11-18 septembre 2023). Saint-Maurice: Santé publique France; 2023. 4 p. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/enquetes-etudes/comment-evolue-la-sante-mentale-des-francais-resultats-de-la-vague-37-de-l-enquete-coviprev>
- [23] Santé publique France. Premiers résultats de l'étude Enabee sur le bien-être et la santé mentale des enfants de 6 à 11 ans en France métropolitaine. Saint-Maurice: Santé publique France; 2023. 10 p. <https://www.santepubliquefrance.fr/etudes-et-enquetes/enabee-etude-nationale-sur-le-bien-etre-des-enfants/documents/premiers-resultats-de-l-etude-enabee-sur-le-bien-etre-et-la-sante-mentale-des-enfants-de-6-a-11-ans-en-france-metropolitaine>
- [24] Vandentorren S, Khirredine I, Estevez M, De Stefano C, Rezzoug D, Oppenchain N, *et al.* Premiers résultats des facteurs associés à la résilience et à la santé mentale des enfants et des adolescents (9-18 ans) lors du premier confinement lié à la Covid-19 en France. *Bull Épidémiol Hebd.* 2021;(Cov_8):2-17. http://beh.santepubliquefrance.fr/beh/2021/Cov_8/2021_Cov_8_1.html
- [25] Araujo-Chaveron L, Doncarli A, Crenn-Hebert C, Demiguel V, Boudet-Berquier J, Barry Y, *et al.* Pregnant women's unmet need to communicate with a health professional during the SARS-CoV-2 pandemic lockdown in France: The Covimater cross-sectional study. *PLoS One.* 2022;17(4):e0266996.
- [26] Bonaldi C, Fouillet A, Sommen C, Lévy-Bruhl D, Paireau J. Monitoring the reproductive number of COVID-19 in France: Comparative estimates from three datasets. *PLoS One.* 2023; 18(10):e0293585.
- [27] Durand J, Fayad M, Feri A, Truong E, Taisne B, Amy E, *et al.* Surveillance virologique en France : impact de la Covid-19 et perspectives. *Bull Épidémiol Hebd.* 2024;(20-21):440-9. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_1.html
- [28] Ministère des Solidarités et de la Santé. Données de la carte de vigilance Covid-19. 2020. <https://www.data.gouv.fr/fr/datasets/donnees-de-la-carte-de-vigilance-covid-19/#/resources>
- [29] Santé publique France. Évolution des indicateurs épidémiologiques et de vaccination chez les enfants et étudiants (au 29 août 2021, semaine 34). Saint-Maurice: Santé publique France; 2021. 9 p. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/enquetes-etudes/evolution-des-indicateurs-epidemiologiques-et-de-vaccination-chez-les-enfants-et-etudiants.-point-au-9-septembre-2021>
- [30] Santé publique France. Diagnostics des passages aux urgences pour Covid-19 (données Oscour®). In: Point Épidémiologique Covid-19 du 03 février 2022. Saint-Maurice: Santé publique France; 2022. pp 9. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-3-fevrier-2022>
- [31] Platon J, Fonteneau L, Hanguhard R, Gagnière B, Gault G, Deschamps G, *et al.* VAC-SI : un système d'information pour le suivi de la couverture vaccinale des vaccins contre la Covid-19. *Bull Épidémiol Hebd.* 2024;(20-21):454-61. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_3.html
- [32] Elliot AJ, Harcourt SE, Hughes HE, Loveridge P, Morbey RA, Smith S, *et al.* The COVID-19 pandemic: A new challenge for syndromic surveillance. *Epidemiol Infect.* 2020;148:e122.
- [33] Ferraro CF, Findlater L, Morbey R, Hughes HE, Harcourt S, Hughes TC, *et al.* Describing the indirect impact of COVID-19 on healthcare utilisation using syndromic surveillance systems. *BMC Public Health.* 2021;21(1):2019.
- [34] Schranz M, Boender TS, Greiner T, Kocher T, Wagner B, Greiner F, *et al.* Changes in emergency department utilisation in Germany before and during different phases of the COVID-19 pandemic, using data from a national surveillance system up to June 2021. *BMC Public Health.* 2023;23(1):799.
- [35] Silk BJ, Scobie HM, Duck WM, Palmer T, Ahmad FB, Binder AM, *et al.* COVID-19 Surveillance After Expiration of the Public Health Emergency Declaration – United States, May 11, 2023. *MMWR Morb Mortal Wkly Rep.* 2023;72(19):523-8.
- [36] Scobie HM, Panaggio M, Binder AM, Gallagher ME, Duck WM, Graff P, *et al.* Correlations and timeliness of COVID-19 surveillance data sources and indicators – United States, October 1, 2020-March 22, 2023. *MMWR Morb Mortal Wkly Rep.* 2023;72(19):529-35.
- [37] Caserio-Schönemann C, Arfaoui A, Coignard B, Le Strat Y, Rolland P, Spacciferri G. Les systèmes d'information à l'épreuve de la Covid-19 : enseignements, nouveaux enjeux et perspectives pour se préparer aux prochaines crises. *Bull Épidémiol Hebd.* 2024;(20-21):488-90. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_7.html

Citer cet article

Fouillet A, Pontais I, Forgeot C, Naud J, Pedrono G, Serra D, *et al.* Surveillance de la pandémie de Covid-19 : contribution et performances du système SurSaUD®. *Bull Épidémiol Hebd.* 2024;(20-21):462-74. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_4.html

SANTÉ PUBLIQUE FRANCE FACE AU DÉFI DE L'OPEN DATA // SANTÉ PUBLIQUE FRANCE RISES TO THE CHALLENGE OF OPEN DATA

Étienne Lucas (etienne.lucas@santepubliquefrance.fr), Aurélie Fouquet, Delphine Jezewski-Serra, Jebraïel Ben Raies, Perrine de Crouy-Chanel, Caroline Alleaume

Santé publique France, Saint-Maurice

Soumis le 18.09.2023 // Date of submission: 09.18.2023

Résumé // Abstract

L'expression « open data » fait référence à un concept dans lequel des données sont mises à disposition de tous, sans restriction de copyright ni de brevet. Il est défini par la disponibilité, la réutilisation des données, et la participation universelle. En santé publique, il permet notamment l'information de tous sur la situation sanitaire, l'enrichissement de la recherche scientifique et l'appui à la gestion ; besoins qui ont été renforcés par la crise de la Covid-19. Cet article vise ainsi à présenter les outils utilisés par Santé publique France pour y répondre, les enjeux, les difficultés rencontrées et les perspectives envisagées.

Santé publique France produit de nombreux indicateurs sanitaires qu'elle met à disposition en toute transparence, ceci en répondant aux obligations réglementaires. Début 2019, l'observatoire cartographique Géodes s'ouvre en ligne avec 300 indicateurs épidémiologiques (900 en juin 2023), facilitant l'exploration et la récupération des données. Les premiers indicateurs liés à la Covid-19 y sont introduits en mars 2020, démultipliant la fréquentation du portail et incitant à recourir à d'autres outils comme la plateforme data.gouv d'Etalab (Direction interministérielle du numérique) puis le tableau de bord de l'agence, InfoCovidFrance. L'utilisation de ces outils pendant la pandémie a permis de placer la France parmi les pays ayant le plus facilité l'accès à leurs données, la classant troisième sur 41 pays évalués sur différents critères.

Le souci de transparence ne fait pas oublier l'impératif de garantir à l'échelle géographique la plus fine la confidentialité et la sécurité des données de santé. Ces éléments et les moyens nécessaires alimentent la réflexion actuelle sur la refonte de l'open data à Santé publique France.

Open data refers to a concept in which data are made available and accessible to all, without copyright or patent restrictions. It is defined by the availability, re-use and redistribution of data, as well as universal participation. In the field of public health, this represents a way of informing society as a whole about the health situation, enriching scientific research, and supporting management – needs that were made more acute by the COVID-19 crisis. The aim of this article is to present the tools used by Santé publique France, the French national public health agency, to meet these needs, as well as the issues at stake, the difficulties encountered and future opportunities under consideration.

Santé publique France produces numerous health indicators, which it makes available in a fully transparent manner while meeting regulatory obligations. In early 2019, the Géodes cartographic observatory was launched online with 300 epidemiological indicators (900 by June 2023), facilitating data exploration and extraction. The first COVID-19 indicators were introduced in March 2020, boosting portal traffic and encouraging the use of other tools such as Etalab's data.gouv platform (French Interministerial Digital Department) and the InfoCovidFrance dashboard. The provision of these tools during the pandemic placed France among the countries that most facilitated access to their data, ranking it third out of 41 countries evaluated on various criteria.

Commitment to transparency does not overlook the need to guarantee the confidentiality and security of health data related to a very fine geographical scale. These considerations, as well as the resources available, contribute to the current discussions on the development of open data at Santé publique France.

Mots-clés : Open data, Portail ouvert, Covid-19, Santé publique France
// Keywords: Open data, Open portal, COVID-19, French National Public Health Agency

Introduction

Dans le cadre de ses missions de surveillance et de prévention des risques sanitaires, Santé publique France, l'Agence nationale de santé publique, produit de nombreux indicateurs renseignant sur

l'état de santé des populations. L'augmentation du nombre d'indicateurs calculés, selon leur disponibilité à différents niveaux géographiques, ainsi que les évolutions en matière d'obligation réglementaire (directive Inspire¹) invitant à la mise à disposition plus large d'informations et de données dans

le champ de la santé ont conduit Santé publique France à se doter en 2019 d'un observatoire cartographique : Géodes⁽¹⁾.

Avec la pandémie de Covid-19 et l'amplification de la production d'indicateurs pour le suivi de l'épidémie, de nouvelles problématiques se sont imposées : réactivité, fréquence et finesse des informations produites, volumétrie, etc. La mise à disposition d'indicateurs sanitaires par Santé publique France a donc dû évoluer avec une mise à jour quotidienne, une diffusion à des échelles fines d'une information sensible, et l'accompagnement de tous les utilisateurs, qu'ils soient issus du grand public, des médias, de la sphère scientifique ou bien des professionnels de santé et de santé publique. L'open data devient en 2021 une partie intégrante de la structuration du programme de travail de Santé publique France via l'enjeu « Numérique en santé publique », un des six enjeux stratégiques réaffirmés en 2023.

Cet article vise à présenter les outils employés par Santé publique France pour répondre aux besoins préexistants et émergents durant la pandémie. Il met en évidence les contraintes et limites des solutions d'open data développées dans un contexte d'urgence en rapport avec l'organisation mise en place (figure). Enfin, il propose des perspectives pour améliorer la stratégie d'open data de Santé publique France afin de mieux satisfaire à l'avenir les attentes d'un public varié.

⁽¹⁾ <https://geodes.santepubliquefrance.fr/>

Matériel et Méthodes

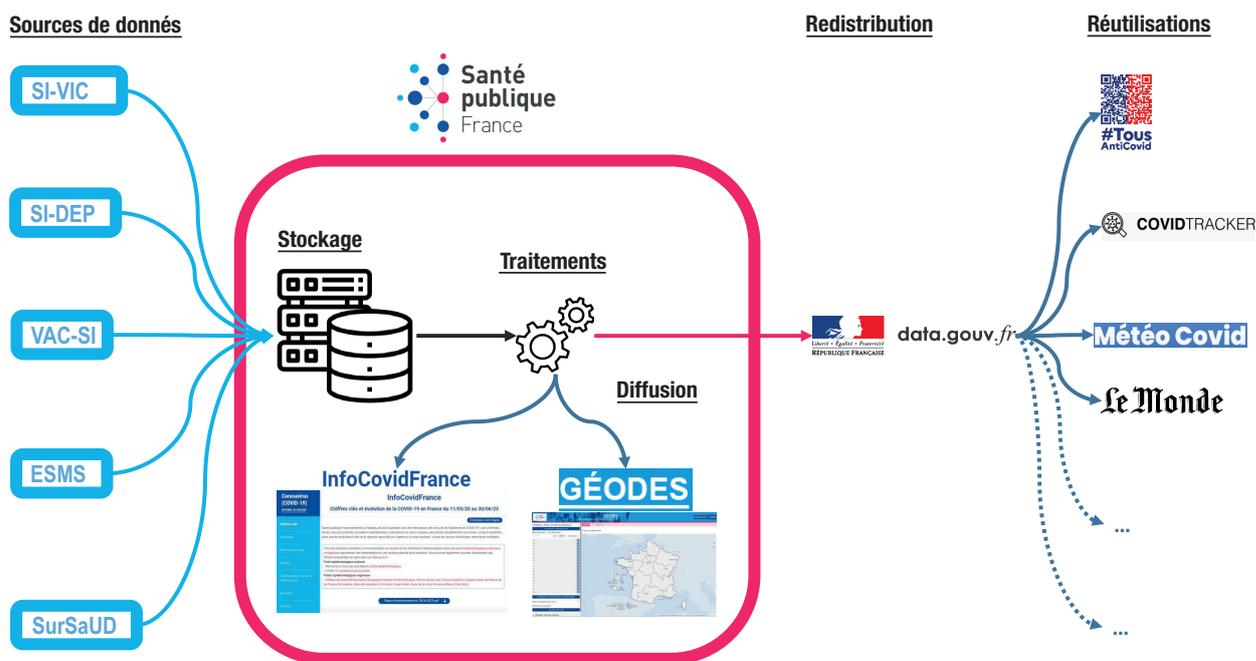
Définition

L'open data fait référence à un concept dans lequel les données sont mises à disposition de manière libre et accessible à tous, sans restriction. Dans ce cadre, les données peuvent être utilisées, réutilisées et redistribuées par n'importe qui, pour n'importe quel but. L'open data est définie par plusieurs principes^{2,3} :

- La disponibilité des données : celles-ci doivent être facilement accessibles par tous ;
- La réutilisation : elle favorise l'innovation, l'efficacité et la transparence en permettant aux individus et aux organisations d'exploiter pleinement le potentiel des données. Par réutilisation, on entend ici les possibilités et fonctionnalités qui vont permettre à l'utilisateur de réutiliser, hors de Géodes, les indicateurs qu'il a consultés, soit sous forme de graphiques, de cartes, ou même en téléchargeant les tables. En effet, « l'utilisation primaire des données de santé renvoie à la collecte de données à l'occasion de la prise en charge des patients, quand l'usage secondaire désigne l'utilisation de ces données pour d'autres finalités comme la recherche et l'innovation, le pilotage ou encore l'amélioration de la qualité des soins. Ainsi, la conception d'algorithmes pour mieux détecter des pathologies, la réalisation d'études pour mesurer les effets en vie réelle de médicaments ou encore l'élaboration de tableaux de bord pour piloter une crise sanitaire s'appuient sur des usages

Figure

Outils d'open data mis en place par Santé publique France pendant la pandémie de Covid-19



SI-VIC : Système d'information pour le suivi des victimes ; SI-DEP : Système d'information de dépistage populationnel ; VAC-SI : Système de suivi des vaccinations ; ESMS : Système de surveillance de la Covid-19 en établissements sociaux et médico-sociaux ; SurSaUD : Surveillance sanitaire des urgences et des décès ; Géodes : Géo données en santé publique.

secondaires des données de santé »⁴. Toutefois il n'est pas toujours garanti que les données obtenues permettent d'aboutir à des résultats probants, fiables, valorisables. En effet, le risque de mauvaise utilisation existe, qu'il relève du simple mésusage ou de l'utilisation à mauvais escient. C'est en rendant disponible une documentation la plus complète et la plus intelligible possible que l'on tend à minimiser ce risque sans pouvoir le supprimer totalement. Afin de tirer le meilleur parti des données, un travail considérable est donc nécessaire pour comprendre leur organisation, la façon dont elles sont collectées, les biais et les limites qu'elles présentent⁵.

- La redistribution : l'open data implique la possibilité de partager les données. Elle peut se faire par le biais de licences ouvertes et encourage ainsi le partage de connaissances et l'innovation. Le concept d'open data est intrinsèquement lié à celui de « *data innovation* »⁶ désignant le fait d'utiliser la donnée pour créer des projets et services à forte valeur ajoutée. Cela implique de mettre les données à disposition de toutes les parties prenantes qui peuvent les réutiliser. Ouvertes et partagées, elles permettent de créer de nouveaux usages innovants.
- La participation universelle : l'open data doit être accessible à tous, sans discrimination ni exclusion. Les données doivent être présentées dans un format compréhensible.

Cadre légal de l'open data

La loi pour une république numérique⁷, adoptée en octobre 2016, constitue un pilier important du cadre légal de l'open data en France. Cette loi introduit le principe de la « donnée d'intérêt général » et établit l'obligation pour les administrations publiques de publier certaines de leurs données dans des formats ouverts et réutilisables. Elle définit également les principes de réutilisation des données publiques, en encourageant la gratuité, la facilité d'accès et la non-discrimination.

La réglementation a été enrichie par d'autres textes et notamment par la directive européenne de 2019⁸ sur les données ouvertes qui a été transposée en droit français en 2021, renforçant ainsi les obligations des administrations publiques en matière de publication de données ouvertes.

Santé publique France, avec sa Direction appui, traitements et analyses des données (Data) créée en 2017, s'organise pour répondre à ces nouvelles obligations.

Résultats

2019 – Géodes

Au début de l'année 2019, Santé publique France inaugure l'observatoire cartographique Géodes (Géo données en santé publique)⁽¹⁾ développé à partir

de la solution Géoclip Air, débutant ainsi le processus de mise en open data des données. Il a été conçu dans le but de rendre accessibles et exploitables les données de santé de l'agence. Le nombre d'indicateurs disponibles est passé de 300 à son ouverture, à plus de 900 en 2023.

Il s'agit d'une plateforme en ligne regroupant et rendant accessible un large éventail d'informations permettant de caractériser l'état de santé des populations en France : des indicateurs de santé issus des données de surveillance, des enquêtes nationales, des bases de données hospitalières, des registres de santé, et des informations sur les déterminants de santé (facteurs qui déterminent l'état de santé des individus). Les thématiques couvertes sont celles de l'agence : maladies infectieuses, maladies chroniques, santé en lien avec l'environnement et santé au travail, comportements de santé, causes de décès, etc. L'objectif de Géodes est de fournir aux chercheurs, professionnels de santé, décideurs et au grand public des informations fiables et actualisées pour la compréhension et l'analyse de différents aspects de la santé.

Géodes offre également des fonctionnalités avancées pour faciliter l'exploration et l'utilisation des données. Les utilisateurs peuvent rechercher, filtrer et télécharger les données en fonction de leurs besoins spécifiques. La plateforme propose des outils de visualisation tels que des cartographies ou des diagrammes en barres pour présenter les données de manière claire et compréhensible, favorisant ainsi l'analyse et la diffusion des informations.

Depuis son ouverture au public, l'outil n'a cessé de s'enrichir de nouvelles thématiques ou des mises à jour de celles existantes. Santé publique France travaille en étroite collaboration avec d'autres acteurs du secteur de la santé : ministère de la Santé, Caisse nationale de l'assurance maladie (Cnam), laboratoires, agences régionales de santé (ARS), sociétés savantes, etc., pour garantir la qualité, la pertinence et la fiabilité des données disponibles sur la plateforme.

Dans un souci de réutilisation des données dans le domaine de la santé publique, les données disponibles sur la plateforme sont mises à disposition sous des licences ouvertes, permettant leur utilisation libre et leur réutilisation par des tiers. Cela contribue à améliorer la compréhension et la gestion de la santé publique.

2020 – L'impact de la Covid-19

La pandémie de Covid-19 a eu un impact significatif sur de nombreux aspects de notre société, y compris le développement de l'open data dans le domaine de la santé, en soulignant l'importance critique de disposer de données de santé fiables, les plus réactives possible et accessibles au grand public. Les informations précises et actualisées sur la propagation du virus, les taux d'infection, les mesures de prévention et les protocoles de traitement sont essentiels pour prendre des décisions

⁽¹⁾ <https://geodes.santepubliquefrance.fr/>

éclairées et mettre en place des politiques de santé efficaces. En conséquence, de nombreux gouvernements et organisations ont intensifié leurs efforts pour rendre les données de santé disponibles au public, encourager la transparence et promouvoir la collaboration entre les différentes parties prenantes.

Les premiers indicateurs Covid-19 ont été disponibles via Géodes en mars 2020. Cette ouverture a permis de garantir la transparence et la réutilisation par le public tout en tenant compte de l'enjeu de protection des données personnelles. Santé publique France a dû trouver l'équilibre dans son offre d'open data en rendant le maximum de données accessibles à l'échelle géographique la plus fine possible, tout en garantissant la non ré-identification des individus. Les indicateurs présentés à une maille géographique fine ne sont ainsi jamais représentés en effectifs et les taux sont uniquement présentés en classes. Il est de ce fait impossible de connaître le nombre de personnes incluses dans le calcul du taux et ainsi de ré-identifier une personne dans un territoire aussi restreint.

La fréquentation du portail, qui restait précédemment limitée et concernait un public averti (médecins, scientifiques, etc.), a alors augmenté de façon exponentielle (de 1 million de visiteurs en 2019 à 15 millions en 2020) et la demande du public s'est élargie. Géodes a vu le nombre d'indicateurs disponibles exploser : environ 200 indicateurs consacrés à la pandémie y sont actualisés quotidiennement entre 2020 et 2023. Toutefois, l'administration du portail s'est révélée contraignante en période de crise. Par ailleurs l'outil Géodes présentait des limites, notamment une navigation qui se complexifie au fur et à mesure que le nombre de thématiques augmente, des possibilités de représentations graphiques relativement réduites, une interface vieillissante et une ergonomie peu intuitive. Enfin, le format des données exportables était trop contraint pour la communauté des *data scientists*⁹ experts en sciences des données, et ne leur permettait pas d'alimenter directement et automatiquement leurs propres outils de visualisation en raison de l'absence d'API⁽²⁾.

La nécessité de passer par d'autres voies s'est imposée : la plateforme data.gouv⁽³⁾ est rapidement apparue comme incontournable, puis la mise à disposition des indicateurs sous forme de tableau de bord (InfoCovidFrance) s'est avérée utile pour compléter par d'autres formes de datavisualisation les informations diffusées sur l'observatoire cartographique Géodes.

data.gouv

Etalab est un département de la Direction interministérielle du numérique (Dinum), chargé de coordonner les efforts de l'administration française pour

rendre les données du secteur public accessibles à tous, dans un format ouvert et réutilisable. Dans cet objectif, Etalab a développé un portail permettant aux administrations, aux entreprises et aux citoyens de publier leurs données : data.gouv. Ce portail est le point focal de l'open data en France depuis sa création en 2011.

Avec l'appui d'Etalab, Santé publique France a ouvert sa page en 2021 sur data.gouv afin de mettre à disposition, de façon quotidienne, les données en lien avec la surveillance de la pandémie de Covid-19. Ce portail a trouvé une forte visibilité auprès de la communauté des *data scientists*, et a permis une excellente diffusion ainsi qu'une large et fréquente réutilisation de ces données.

La mise à disposition sur data.gouv a ainsi permis à de nombreux acteurs de valoriser la donnée à travers d'autres outils de restitution. On compte ainsi plus de 250 réutilisations des données de l'agence (144 réutilisations pour les données d'hospitalisations, 44 pour l'activité de dépistage, 52 pour l'activité SOS Médecins et des urgences, 46 pour le suivi des campagnes de dépistage, etc.). Ces réutilisations prennent plusieurs formes, comme des tableaux de bord, des API ou des data-visualisations. Elles sont produites par un éventail d'acteurs, qu'ils soient publics, privés ou citoyens. On peut citer : Covidtraker⁽⁴⁾, Météo-Covid⁽⁵⁾, les Décodeurs du Monde⁽⁶⁾, TousAntiCovid, le Tableau de bord du gouvernement, Architecture&Performance⁽⁷⁾.

L'utilisation de data.gouv par Santé publique France pour publier les données de surveillance de la Covid-19 a été efficace. La France a été remarquée pendant la pandémie comme l'un des pays ayant le plus facilité l'accès à ses données¹⁰. Elle est classée troisième sur 41 pays évalués par différents critères parmi lesquels : description des données, date de publication et rythme de mise à jour, forums et formulaires de contact, documentation et tutoriels, visualisation et statistiques, partage des données, etc. La France, l'Autriche et le Portugal sont les 3 pays qui apparaissent le plus fréquemment, pour chaque critère, parmi les 5 pays en tête de ce classement.

InfoCovidFrance

Ainsi, la crise sanitaire a vu fortement augmenter l'offre de tableaux de bord accessibles au public présentant une description de la situation sanitaire au prime des indicateurs d'infections, d'hospitalisation, de décès, ainsi que de vaccination liés à la pandémie. Dans un contexte de forte médiatisation, ces indicateurs sont rapidement devenus des éléments essentiels de communication des médias, des établissements, ainsi que des autorités de santé.

⁽²⁾ L'API ou Application Programming Interface (Interface de programmation d'application) est l'ensemble normalisé de classes, de méthodes, de fonctions, de constantes – des données – par lequel un logiciel offre des services à un autre logiciel qui les consomme pour alimenter ses propres fonctionnalités.

⁽³⁾ <https://www.data.gouv.fr/fr/organizations/sante-publique-france/>

⁽⁴⁾ <https://covidtracker.fr/>

⁽⁵⁾ <https://www.meteo-covid.com/>

⁽⁶⁾ https://www.lemonde.fr/les-decodeurs/article/2022/11/29/covid-19-les-chiffres-de-l-epidemie-en-france-et-dans-le-monde-en-cartes-et-en-graphiques_6038751_4355777.html

⁽⁷⁾ https://www.architecture-performance.fr/ap_blog/tests-incidence-hospitalisations-reanimations-et-deces-du-au-covid-en-france-par-departements-et-par-regions/

Les tableaux de bord Covid-19 destinés au grand public sont ainsi plus largement démocratisés que pour tout autre problème de santé¹¹. Si la mise à disposition des indicateurs de suivi de l'épidémie de Covid-19 a pu dès 2020 être réalisée via Géodes puis par data.gouv, permettant l'émergence de ces nouvelles plateformes de data-visualisation, Santé publique France développe en 2021 son propre tableau de bord. InfoCovidFrance⁽⁸⁾, accessible depuis le site Internet de l'agence, permet de publier pour la première fois de nouveaux indicateurs sur la pandémie issues du *contact-tracing*, sur l'activité de la Réserve sanitaire, de distribution des vaccins ou encore sur les résultats d'enquête sur l'adhésion aux comportements de prévention. Plus qu'un outil de valorisation de l'activité de Santé publique France, InfoCovidFrance vise à être pédagogique pour être accessible à un plus grand nombre. Les visuels varient entre des graphiques illustrant la donnée, des courbes de tendances, des formats chiffres clés permettant d'interpréter en un coup d'œil l'indicateur, et enfin des infographies illustrant simplement l'histoire racontée par la donnée¹². Ainsi le tableau de bord offre un outil très visuel, interactif, favorisant l'appropriation des informations, ainsi que des entrées à la fois thématiques, populationnelles et géographiques. Enfin, InfoCovidFrance propose également une version anglaise visant à augmenter la visibilité et à faciliter le partage d'information dans le cadre de collaborations internationales.

Accompagnement des utilisateurs

Dans un contexte de réutilisation des données publiées sur différents supports (dont les réseaux sociaux) et d'interprétation des résultats par un public de plus en plus expert, l'agence a très vite accompagné les utilisateurs dans la compréhension de la donnée.

Les données fournies par Santé publique France avaient pour rôle d'informer en totale transparence la population de l'évolution de l'épidémie. La quantité d'informations a soulevé de nombreuses interrogations et l'agence a dû répondre aux nombreuses sollicitations reçues par le biais des formulaires de contact ou tweets émanant de citoyens, de scientifiques ou de journalistes. Ce sont en tout près de 4 000 formulaires qui ont été remplis et envoyés via Géodes entre mars 2020 et juillet 2022, soit 4 par jour en moyenne. En comparaison, entre juillet 2022 et juillet 2023 nous avons reçu seulement 432 formulaires de contact.

Les outils de visualisation ont été un mode de consommation privilégié des données par le grand public. On peut ainsi penser aux différents graphiques produits par des data-journalistes qui sont facilement consultables sur les réseaux sociaux, comme sur Twitter (aujourd'hui X). La mise à disposition de la donnée

sur data.gouv et sa consommation sur les différents outils de restitution ont suscité un intérêt plus vif du grand public. Certains utilisateurs développaient une réelle expertise et se tournaient naturellement vers le producteur de la donnée (Santé publique France). Ainsi, l'onglet discussion de la page data.gouv a été fortement sollicité pendant cette période.

Pour les publications sur InfoCovidFrance, data.gouv et Géodes, chaque jeu de données et chaque indicateur étaient accompagnés par une description précise dans l'onglet dédié. Parfois, des notes méthodologiques étaient diffusées. C'est le cas pour la page SI-DEP, et sa fiche méthodologique téléchargeable¹³ pour préciser les changements qui ont eu lieu sur les indicateurs issus de la surveillance virologique à compter de mai 2022. Cette dernière a été téléchargée plus de 600 fois. Des moments d'accompagnement pédagogique étaient planifiés à l'attention des principaux utilisateurs (journalistes et *data scientists* en particulier) avant la publication de nouveaux indicateurs ou tout autre évolution de l'offre disponible. Ces moments fournissaient l'occasion de présenter les nouvelles publications, de répondre aux questions des utilisateurs, de prendre en considération leurs commentaires et remarques pour améliorer la qualité de la publication officielle.

Discussion

Optimisation des outils

La multiplication des plateformes open data complique la lisibilité pour les citoyens, et les différences méthodologiques de construction entre les indicateurs produits par les différents acteurs peuvent susciter des questionnements voire même de la défiance, comme cela a pu être constaté dans les messages reçus via les formulaires de contact Géodes. Forte de ce constat, Santé publique France s'organise pour engager une refonte de ses outils d'open data venant pallier les limites de Géodes en vue d'améliorer l'expérience utilisateur, de proposer une palette plus large de représentations graphiques, de renforcer la lisibilité et l'intelligibilité des informations mises à disposition et d'en faciliter l'administration.

Protection des données personnelles

Les données de santé sont extrêmement sensibles, contenant des informations personnelles et confidentielles. Lorsque ces données sont rendues accessibles au public dans le cadre de l'open data, il est impératif de garantir leur confidentialité et leur sécurité, notamment en empêchant la ré-identification par une anonymisation⁽⁹⁾ rigoureuse avant leur publication.

En assurant la protection des données de santé et en se préoccupant de leur qualité, l'open data

⁽⁸⁾ <https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolution-de-la-covid-19-en-france-et-dans-le-monde>

⁽⁹⁾ L'anonymisation de données consiste à modifier le contenu ou la structure de ces données afin de rendre très difficile ou impossible la ré-identification des personnes ou des entités concernées.

est devenu une ressource précieuse pour la santé publique. Les chercheurs, les professionnels de santé et les décideurs politiques peuvent exploiter ces données pour améliorer les soins, identifier les tendances épidémiologiques et prendre des décisions de gestion et de prévention éclairées basées sur des preuves.

Data.gouv étant un point focal de l'open data en France, Santé publique France ne pourra pas, à l'avenir, se priver de sa visibilité. Une réflexion doit cependant être menée pour savoir ce qui doit y être diffusé. Les données Covid-19 devaient y être déposées, compte tenu du rythme quotidien de mise à jour et d'un fort intérêt pour ces données dans l'espace médiatique. La question est ouverte concernant les autres données produites par l'agence (issues des systèmes de surveillance ou d'enquêtes) et leur place sur un portail généraliste d'open data. La complexité de certains indicateurs et de leurs méthodes de construction pose également la question du public auquel l'open data s'adresse, et du niveau de littératie en santé des utilisateurs, pour les rendre accessibles et compréhensibles par le plus grand nombre.

Refonte du système pour une homogénéisation des pratiques

Sortie du contexte de crise de la pandémie, à moyens et ressources diminués, Santé publique France ne peut pertinemment pas gérer deux outils de diffusion d'indicateurs en open data en complément de la mise à disposition de données sur data.gouv. Une réflexion actuellement en cours va probablement mener vers une disparition des outils Géodes et InfoCovidFrance au profit d'un environnement rassemblant les avantages des deux outils avec une qualité de service équivalente, en différenciant deux profils d'utilisateurs. Un premier profil « grand public » pourra naviguer intuitivement pour visualiser les données et indicateurs directement dans l'application avec une documentation vulgarisée. Les experts préféreront consulter une documentation scientifique et interroger les bases de données via les API. L'objectif de cette refonte sera d'avoir à disposition un outil incluant un tableau de bord avec des chiffres clés sur les surveillances sanitaires dont l'actualité mérite une mise en avant (par exemple : la grippe en période hivernale, la surveillance des décès suite à des vagues de chaleurs, des résultats d'enquêtes...), des indicateurs couvrant l'ensemble des champs de surveillance de l'agence.

Conclusion

La nécessité urgente, induite par la pandémie de Covid-19, de partager des informations en masse et de collaborer à l'échelle mondiale a incité de nombreux pays et organisations à adopter des approches plus ouvertes en matière de données de santé. De nouvelles initiatives et plateformes ont vu le jour, permettant aux chercheurs et aux décideurs politiques d'accéder à des ensembles de données précieux pour mieux comprendre la propagation de

l'épidémie, évaluer l'efficacité des mesures de santé publique et développer de nouveaux traitements. En France la stratégie d'open data est impulsée au niveau national depuis les années 2010. Il existe des normes d'interopérabilité qui se sont structurées, permettant de bénéficier d'outils de mise à disposition des indicateurs produits par l'agence et d'accéder à des jeux de données d'autres institutions. La diversité des systèmes de santé à travers le monde et les variations dans la collecte et la normalisation des données ont rendu difficile la mise en place d'un cadre universel d'open data en santé dans le contexte d'urgence induit par la pandémie. Les différences de terminologie, de format et de méthodologie de collecte des données engendrent des défis lors de l'agrégation, de l'analyse et du croisement des informations issues de sources très diverses. Pour maximiser l'impact de l'open data dans le domaine de la santé et associer les données à l'action, il sera crucial de développer des normes et des protocoles communs pour la collecte, la mise en qualité, le traitement et le partage des données^{4,6}. ■

Remerciements

Nous remercions tous les acteurs de Santé publique France qui ont participé à la mise en place, la collecte, l'organisation des flux de données, aux tests, à la communication et la maintenance des différents outils d'open data : Johnny Platon, Benjamin Taisne, Édouard Chatignoux, Camille Pelat, Cécile Sommen, Céline Caserio-Schönemann, Jérôme Guillevic, Andrea Guajardo Villar, Yann Le Strat, Guillaume Spacciferri, Alima Marie-Malikite, Laëtitia Huiart, Stéphane Nardy, Adel Arfaoui, Paul-Henri Lampe, Sophie Mallejac, Minh-Canh Quan, Michel Slimane, Céline Zaccarini, Anne Robion, Marie Hamsany, Aurélien Dousseron, Anouk Tabai.

Nous remercions également nos partenaires et leurs équipes : EtaLab, TousAntiCovid, CirilGroup et DigDash.

Liens d'intérêts

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

- [1] Ministère du Partenariat avec les territoires et la Décentralisation. La directive européenne Inspire. 2019. <https://www.ecologie.gouv.fr/politiques-publiques/directive-europeenne-inspire>
- [2] James L. Defining Open Data. 2013. <https://blog.okfn.org/2013/10/03/defining-open-data/>
- [3] Commission nationale informatique et libertés. Guide pratique de la publication en ligne et de la réutilisation des données publiques (« open data »). Paris: Cnil; 2019. 28 p. <https://www.cada.fr/lacada/guide-pratique-de-la-publication-en-ligne-et-de-la-reutilisation-des-donnees-publiques>
- [4] Marchand-Arvier J, Allasonnière S, Hoang A, Jannot AS. Fédérer les acteurs de l'écosystème pour libérer l'utilisation secondaire des données de santé. Paris: ministère de la Santé et de la Prévention; 2023. 150 p. https://sante.gouv.fr/IMG/pdf/rapport_donnees_de_sante.pdf
- [5] Health Data Hub. Innover avec l'ensemble des acteurs. 2024. <https://www.health-data-hub.fr/innover-avec-lensemble-des-acteurs>
- [6] Ministère de la Santé et de la Prévention. 18. Développer la recherche en santé numérique et en particulier l'utilisation secondaire des données de santé. In: Feuille de route du

numérique en santé 2023-2027. Paris: ministère de la Santé et de la Prévention; 2023. p. 39. <https://sante.gouv.fr/actualites/actualites-du-ministere/article/lancement-de-la-feuille-de-route-du-numerique-en-sante-2023-2027>

[7] LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000033202746/>

[8] Union européenne. Directive (UE) 2019/1024 du Parlement européen et du Conseil du 20 juin 2019 concernant les données ouvertes et la réutilisation des informations du secteur public (refonte). <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32019L1024>

[9] Concepteurs d'avenir. Data scientist. 2024. <https://www.concepteursdavenir.fr/fiche-metier/data-scientist-0>

[10] Nikiforova A, McBride K. Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*. 2021;(58):101539.

[11] Dasgupta N, Kapadia F. The future of the public health data dashboard. *Am J Public Health*. 2022;112(6):886-8.

[12] Chau D, Parra J, Santos MG, Bastías MJ, Kim R, Handley MA. Community engagement in the development of health-related data visualizations: A scoping review. *J Am Med Inform Assoc*. 2024;31(2):479-87.

[13] data.gouv. Données de laboratoires pour le dépistage (à compter du 18/05/2022) – SI-DEP. 2023. <https://www.data.gouv.fr/fr/datasets/donnees-de-laboratoires-pour-le-depistage-a-compter-du-18-05-2022-si-dep/>

Citer cet article

Lucas É, Fouquet A, Jezewski-Serra D, Ben Raies J, de Crouy-Chanel P, Alleaume C. Santé publique France face au défi de l'open data. *Bull Épidémiol Hebd*. 2024;(20-21):475-81. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_5.html

> ARTICLE // Article

SURVESMS : UN DISPOSITIF DE SURVEILLANCE DE LA COVID-19 EN ÉTABLISSEMENTS SOCIAUX ET MÉDICO-SOCIAUX

// SURVESMS: A SYSTEM FOR MONITORING COVID-19 IN LONG-TERM CARE FACILITIES

Côme Daniau¹ (come.daniau@santepubliquefrance.fr), Jérôme Naud¹, Sabrina Tessier², Éline Hassan³

¹ Santé publique France, Saint-Maurice

² Santé publique France – Bourgogne-Franche-Comté, Dijon

³ Santé publique France – Antilles, Gourbeyre

Soumis le 28.02.2024 // Date of submission: 02.28.2024

Résumé // Abstract

Lors de la phase de diffusion du virus SARS-CoV-2, un des enjeux majeurs de la surveillance épidémiologique a été de disposer de données actualisées et exhaustives sur les personnes atteintes de Covid-19 dans les établissements sociaux et médico-sociaux (ESMS). Cet article décrit la mise en place du dispositif de surveillance de la Covid-19 en ESMS (SurvESMS) et son adaptation pour répondre aux besoins des acteurs de la surveillance et des décideurs, ainsi qu'à l'évolution des pratiques face à l'acquisition progressive des connaissances sur la Covid-19. Le développement de ce dispositif de surveillance met en lumière la nécessité de se doter de plateformes informatiques performantes et agiles, permettant d'accéder aux données de santé des résidents et d'être en capacité de développer rapidement des systèmes de surveillance en situation d'urgence.

During the outbreak of the SARS-CoV-2 virus, one of the major challenges for epidemiological surveillance was to obtain up-to-date, exhaustive data on people with COVID-19 in long-term care facilities. This article describes the implementation of the COVID-19 surveillance system in long-term care facilities (SurvESMS). It also describes how the system was adapted to meet the needs of people involved in surveillance and decision-makers, as well as the changes in practices resulting from the gradual acquisition of knowledge about COVID-19. The development of this surveillance system highlights new requirements for agile IT platforms that provide access to residents' health data and that can be used to develop surveillance systems rapidly in the event of an outbreak.

Mots-clés : Dispositif de surveillance, Covid-19, Établissements sociaux et médico-sociaux, Solution informatique
// Keywords: Surveillance system, COVID-19, Residential care, IT solution

Introduction

Lors de la pandémie de Covid-19, il a été rapidement mis en évidence que certaines populations étaient plus à risque de formes graves de la maladie, en particulier les personnes âgées¹⁻³. Les acteurs de terrain ont rapporté très tôt que l'âge était l'un des principaux facteurs de risque associés à la mortalité liée à la Covid-19 ; le taux de mortalité étant 60 fois plus élevé chez les personnes âgées de plus de 80 ans par rapport à une population âgée de moins de 50 ans⁴. En particulier, des études ont rapporté de nombreuses situations épidémiques dans les établissements de long séjour pour personnes âgées avec des taux de mortalité très élevés⁵⁻⁸.

Lors de la phase de diffusion du virus, disposer de données réactives et exhaustives sur les personnes atteintes de Covid-19 dans les établissements sociaux et médico-sociaux (ESMS) est devenu un enjeu majeur de la surveillance épidémiologique. Parmi l'ensemble des dispositifs de surveillance multisources mis en œuvre pour répondre à cette infection émergente⁹, Santé publique France a mis en œuvre une surveillance de la Covid-19 spécifique de cette population à risque.

L'objectif de cet article est de décrire la mise en place du dispositif de surveillance de la Covid-19 en ESMS (SurvESMS). Il détaille les données recueillies, leurs traitements, les indicateurs produits et la solution informatique proposée avec ses avantages et ses limites. Il vise également à décrire les évolutions qui ont été apportées au dispositif de surveillance. Les résultats de surveillance issus de ce dispositif ne sont pas présentés dans cet article.

Déploiement en urgence d'un dispositif de surveillance de la Covid-19 en ESMS

À la suite de l'évolution de la situation épidémiologique et à la parution du MINSANTE/CORRUS n° 2020_46 du 27 mars 2020 du ministère de la Santé¹⁰, Santé publique France a développé en urgence un dispositif de surveillance des cas de Covid-19 en ESMS afin de répondre aux demandes du ministère de la Santé. La Direction générale de la santé (DGS) sollicitait une remontée quotidienne du nombre de cas et de décès liés à la Covid-19 dans les ESMS pendant au moins les périodes d'intense circulation du virus.

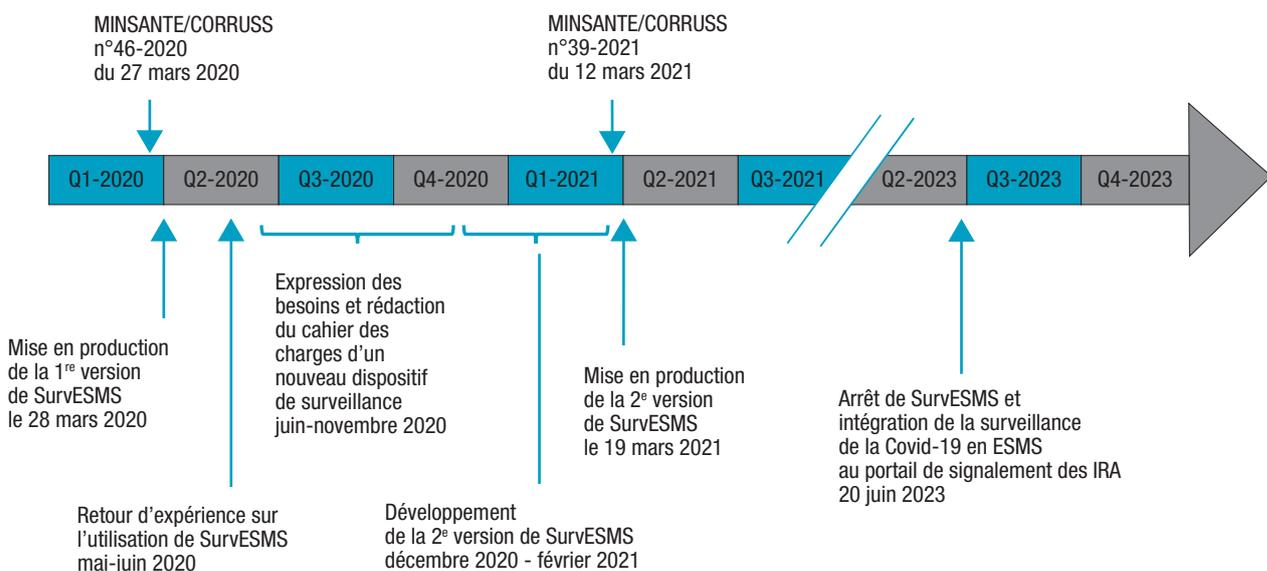
L'objectif de ce dispositif était d'assurer la détection rapide des personnes infectées par le SARS-CoV-2 parmi les résidents et le personnel des ESMS afin de mettre en place des mesures de gestion dans les meilleurs délais et d'assurer le suivi du nombre de cas et de décès liés à la Covid-19 en temps réel. Le dispositif visait l'élaboration de rapports quotidiens de la situation épidémique.

Santé publique France ayant été informée en amont de la parution du MINSANTE/CORRUS, le dispositif de surveillance SurvESMS a été mis à disposition des ESMS le 28 mars 2020 pour le signalement, sur la base du volontariat, des cas et des décès liés à la Covid-19 (figure 1).

La surveillance a concerné un éventail très large de catégorie d'ESMS. Les ESMS ciblés par la surveillance étaient les établissements d'hébergement pour personnes âgées (Ehpa), notamment pour personnes âgées dépendantes (Ehpad), les établissements d'hébergement pour personnes en situation de handicap (HPH), les établissements d'aide sociale à l'enfance et d'autres ESMS comme les centres

Figure 1

Chronologie de la mise en place du dispositif de surveillance de la Covid-19 en établissements sociaux et médico-sociaux (SurvESMS), France



IRA : infections respiratoires aiguës.

de soins et d'accompagnement et de prévention en addictologie (CSAPA) avec hébergement, et, plus largement, les établissements d'hébergement pour personnes précaires. Les établissements sans place d'hébergement ne faisaient pas partie du champ de la surveillance.

Outre les ESMS sollicités pour signaler dans SurvESMS les cas et les décès de Covid-19 survenus dans leur établissement, le dispositif impliquait les agences régionales de santé (ARS) et les centres régionaux d'appui pour la prévention des infections associées aux soins (CPIas) pour l'évaluation de la situation et l'appui à la mise en place des mesures de gestion. Les cellules régionales de Santé publique France intervenaient en appui à la gestion en apportant leur expertise épidémiologique. Santé publique France produisait des indicateurs communiqués au niveau national au ministère de tutelle et aux acteurs régionaux pour l'aide à la gestion (figure 2).

Le critère de signalement initial d'un épisode de cas de Covid-19 était la survenue du premier cas possible ou confirmé de Covid-19 chez les résidents ou le personnel¹¹. La distinction des cas de Covid-19 possibles et confirmés avait son importance en début de pandémie lorsque la confirmation de l'infection par le résultat positif d'un test RT-PCR ou d'un test antigénique était rarement obtenue. Les cas possibles de Covid-19 étaient le plus souvent définis par la présence de signes et symptômes respiratoires ou tout autre tableau clinique compatible avec la Covid-19.

L'épisode était ensuite actualisé par le signalement quotidien des nouveaux cas de Covid-19 et des nouveaux décès attribuables à la Covid-19, y compris en l'absence de nouveau cas. Enfin, 14 jours après la survenue du dernier cas dans l'établissement, l'ESMS effectuait un signalement afin de clôturer

l'épisode. Chaque étape du signalement faisait l'objet d'un formulaire électronique complété par l'établissement et transmis après enregistrement à Santé publique France et conservé dans une base de données sécurisée.

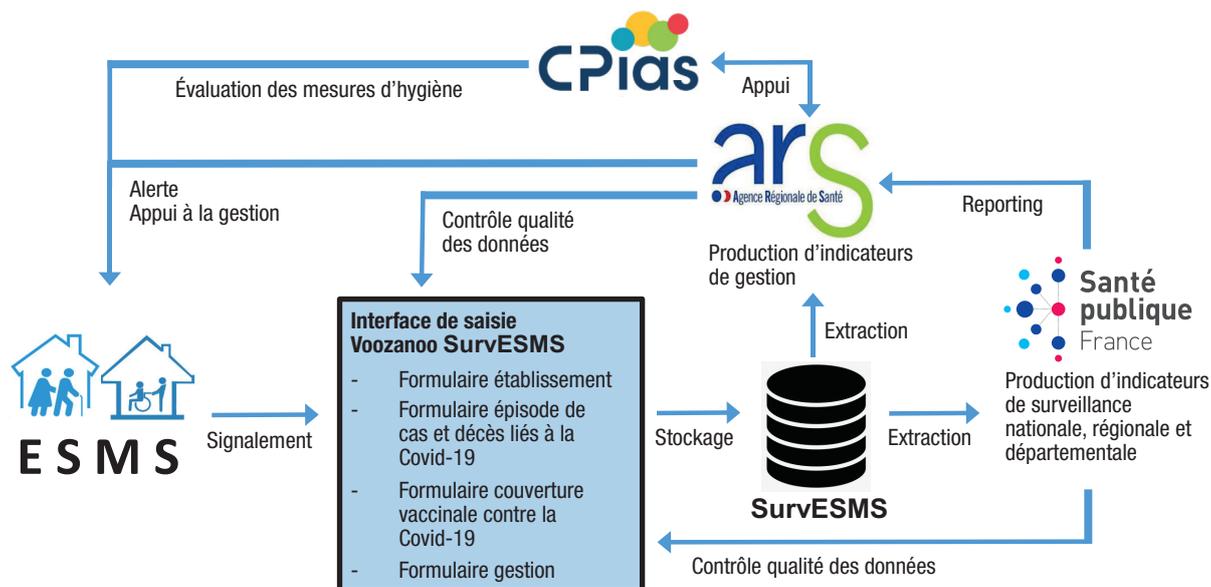
Les données de santé recueillies, exclusivement agrégées par établissement, portaient sur le nombre de cas possibles ou confirmés de Covid-19 chez les résidents, dont ceux hospitalisés, et le nombre de décès liés au Covid-19 chez les résidents dont ceux décédés à l'hôpital. Concernant le personnel des ESMS, le nombre de cas possibles ou confirmés de Covid-19, dont ceux hospitalisés, étaient recueillis. Le formulaire recueillait également le nombre de tests pour la recherche du SARS-CoV-2 réalisés chez les résidents et les membres du personnel de l'établissement.

Des métadonnées étaient collectées et portaient sur la date de signalement, la personne déclarant le signalement (nom, fonction, téléphone, fax, email) et les caractéristiques de l'établissement (raison sociale, code postal, région, département, commune, numéro du Fichier national des établissements sanitaires et sociaux (Finess) géographique, rattachement à un établissement de santé, catégorie d'établissement, nombre de résidents et de personnes employées dans l'établissement au moment du signalement).

Des informations relatives aux mesures mises en œuvre pour la gestion de l'épisode de Covid-19 étaient également recueillies : la mise en place d'une cellule de crise pour la gestion de l'épisode, l'application des recommandations de la Société française d'hygiène hospitalière (SF2H) pour la prévention de la transmission croisée des micro-organismes¹²⁻¹⁴, l'éventuelle suspension des admissions dans l'établissement, la réalisation du bionettoyage adapté

Figure 2

Rôle des acteurs de la surveillance et de la gestion des épisodes de Covid-19 en établissements sociaux et médico-sociaux (ESMS)



ARS : agences régionales de santé ; CPIas : centres régionaux d'appui pour la prévention des infections associées aux soins.

à la Covid-19, l'accès à des équipements de protection individuelle (EPI) en quantité suffisante, la mise en œuvre d'une information des familles et du personnel. Enfin, l'ESMS pouvait formuler une demande de soutien extérieur de l'ARS.

Les données ont permis de produire des indicateurs, stratifiés selon la catégorie d'établissement, sous forme de graphiques et de tableaux relatifs au nombre d'épisodes de cas de Covid-19 totaux, en cours ou clôturés, au nombre de cas de Covid-19 possibles et confirmés chez les résidents dont ceux hospitalisés, au nombre de décès liés à la Covid-19 chez les résidents dont ceux décédés à l'hôpital et, enfin, au nombre de cas de Covid-19 possibles et confirmés chez le personnel des ESMS. Tous les indicateurs étaient produits quotidiennement à l'aide du logiciel Stata®.

Choix de la solution informatique dans sa version initiale : avantages et limites

Au début de la pandémie de Covid-19, Santé publique France ne disposait pas de système de surveillance sanitaire spécifique aux ESMS en capacité d'intégrer la surveillance de la Covid-19. En conséquence, une solution informatique a été proposée pour mettre en œuvre le dispositif de surveillance de la Covid-19 en ESMS.

Depuis 2019, le signalement de cas groupés d'infections respiratoires aiguës (IRA) et de gastro-entérites aiguës (GEA) en Ehpad s'effectue via le portail de signalement des événements sanitaires indésirables du ministère de la Santé¹⁵. Or, les critères de signalement des cas groupés d'IRA en Ehpad différaient de ceux nécessaires en début de pandémie pour la surveillance des épisodes de cas de Covid-19¹⁰. Alors que le critère de signalement initial d'un épisode de cas de Covid-19 était la survenue d'un premier cas possible ou confirmé de Covid-19 chez les résidents ou le personnel, le critère de signalement d'un épisode de cas groupés d'IRA se définissait par l'identification d'au moins 5 cas d'IRA dans un délai de 4 jours. De plus, la fréquence de signalement quotidienne attendue pour la surveillance de la Covid-19, en particulier pendant les périodes de pic épidémique, était incompatible avec les modalités de signalement des épisodes d'IRA qui prévoit, pour chaque épisode, seulement deux signalements, l'un initial et l'autre à la clôture de l'épisode. Le portail de signalement des événements sanitaires indésirables du ministère de la Santé était donc inadapté.

L'adaptation du portail du signalement des événements sanitaires indésirables pour le signalement des cas de Covid-19 n'a pas été possible dans des délais compatibles avec la mise en place en urgence du dispositif de surveillance. En outre, il a semblé qu'une prestation pour le développement d'une nouvelle solution informatique n'était pas adaptée car trop longue à mettre en œuvre dans un contexte

d'urgence. Le choix de la solution informatique s'est donc porté sur la plateforme Voozoo 3, plateforme développée par la société EpiConcept et utilisée à Santé publique France. Voozoo présente une grande souplesse et permet, via un éditeur de contenu, de créer des formulaires pour le recueil de données et de gérer les droits des utilisateurs et les flux de données. En utilisant cette plateforme, Santé publique France était en capacité de développer un outil informatique dans un délai très court.

La plateforme Voozoo 3 a permis de proposer une solution informatique pour la surveillance des cas et des décès liés à la Covid-19 répondant aux principaux objectifs de surveillance du MINSANTE/CORRUSS¹⁰. Les données recueillies quotidiennement par les ESMS sur les épisodes de Covid-19 ont permis de produire les principaux indicateurs de la surveillance nationale et régionale via les bulletins quotidiens, les points épidémiologiques (PE) hebdomadaires et les tableaux de bords, ainsi qu'une valorisation de ces résultats¹⁶⁻²⁰.

Cependant, les utilisateurs de l'outil informatique – les ESMS dont les Ehpad, ayant déjà une culture du signalement, et les acteurs régionaux, les ARS et les cellules régionales de Santé publique France – ont rapidement fait remonter des difficultés et des faiblesses de l'outil informatique. Un retour d'expérience a été mené par Santé publique France avec les utilisateurs dans les mois qui ont suivi la mise en production de l'outil informatique (figure 1).

Tout d'abord, l'outil informatique présentait un problème d'ergonomie. Les utilisateurs étaient confrontés à une interface de saisie vieillissante, peu intuitive, sans contrôle des données saisies et avec l'obligation à chaque signalement de saisir à nouveau des données administratives sur l'établissement et les informations sur le déclarant du signalement.

De plus, l'outil ne permettait pas aux établissements signalants d'accéder aux données transmises, afin de pouvoir les modifier ce qui posait des problèmes de qualité des données. Les corrections des erreurs de saisie ou des doublons de cas et de décès liés à la Covid-19 étaient effectuées a posteriori par les cellules régionales de Santé publique France, avec pour conséquence une surcharge de travail et générant des résultats difficiles à communiquer.

En outre, les indicateurs de la surveillance n'étaient pas mis à disposition des acteurs régionaux de manière harmonisée afin de répondre aux besoins d'appui à la gestion des épisodes de cas (figure 2). Des cellules régionales produisaient des indicateurs régionaux pour les ARS pour pallier le manque de retour d'information régional. L'outil a été pensé principalement pour la surveillance épidémiologique de la Covid-19 au niveau national, sans répondre aux besoins régionaux et locaux.

Enfin, les critères de signalement pour la surveillance de la Covid-19 en ESMS ont évolué au cours de la pandémie, notamment en matière de fréquence de déclaration des cas et de diagnostic des cas,

exclusivement cliniques au début de la pandémie puis, progressivement, confirmés biologiquement. De nouveaux indicateurs étaient nécessaires pour caractériser la dynamique épidémique.

Évolution du dispositif de surveillance SurvESMS

Tout en poursuivant la surveillance de la Covid-19 en ESMS, Santé publique France a amélioré le dispositif. Une évaluation des besoins a été menée au sein d'un groupe de travail dans lequel les cellules régionales de Santé publique France ont porté les attentes régionales. Des évolutions ont été apportées au dispositif de surveillance SurvESMS, dont la nouvelle version a été mise à disposition des ESMS le 19 mars 2021 (figure 1). À ce stade de la pandémie, les établissements pour personnes âgées restaient les principaux déclarants.

Des contrôles de cohérence des données entre les formulaires de l'ensemble des signalements de cas et de décès d'un même épisode ont été mis en œuvre de manière à améliorer la qualité des données. La gestion des comptes des utilisateurs dans les ESMS a permis aux utilisateurs d'accéder à l'historique de leurs données pour éventuellement les modifier. Une deuxième application dédiée à la création des comptes a été développée sous LimeSurvey (LimeSurvey GmbH) pour fluidifier la gestion des demandes des nouveaux utilisateurs déclarants.

Compte tenu de l'évolution de la surveillance de la Covid-19, des modifications ont été apportées aux critères et à la fréquence de signalement. D'une part, avec la généralisation des tests diagnostiques, seuls les cas confirmés de Covid-19 étaient signalés. D'autre part, la mise à jour quotidienne du signalement en l'absence de nouveau cas n'était plus nécessaire. Les signalements étaient recommandés uniquement lors de la survenue de nouveaux cas ou de nouveaux décès dans l'établissement.

L'ensemble des champs relatifs à la gestion a été dissocié de la surveillance sanitaire. Un nouvel outil dédié à l'aide à la gestion d'une épidémie de Covid-19 en ESMS a été développé par la mission nationale de surveillance et prévention de la résistance aux antibiotiques et des infections associées aux soins en soins de ville et en secteur médico-social (MN-Primo). Seul le champ relatif à la demande d'un soutien extérieur de l'ARS pour faire face à l'épisode en cours a été conservé et pouvait être actualisé. En outre, un formulaire spécifique sur la surveillance de la couverture vaccinale contre la Covid-19 en ESMS a été ajouté à l'outil informatique pour compléter le dispositif de surveillance SurvESMS.

De nouveaux indicateurs de surveillance ont été produits : le nombre de foyers épidémiques comptant au moins trois cas confirmés, le taux d'attaque, le taux d'hospitalisation, la létalité chez les résidents et le personnel, la répartition des épisodes stratifiés selon le nombre de cas, la durée des épisodes. À l'inverse, certains indicateurs ont été abandonnés

en raison de l'arrêt de leur recueil, comme le nombre de cas possibles de Covid-19 et le nombre de tests pour la recherche du SARS-CoV-2.

La restitution des indicateurs de la surveillance auprès des acteurs de terrain et notamment des ARS a été conçue de manière intégrée et fondée sur l'exploitation automatisée des données. Des programmes de contrôle de la qualité des données, d'analyses et de restitution des résultats ont été élaborés sous R Markdown. Des rapports de résultats au format html étaient produits quotidiennement et de manière automatisée aux niveaux national, régional et départemental. Les résultats présentaient les indicateurs de la surveillance actualisés, cohérents entre les différents niveaux géographiques et stratifiés par catégorie d'établissement. Des courbes épidémiques à différentes échelles géographiques présentaient l'évolution du nombre hebdomadaire d'épisodes, de foyers épidémiques, de cas et de décès liés à la Covid-19. De plus, des fichiers au format csv pour chaque région donnaient des informations détaillées sur les caractéristiques des épisodes en cours et clôturés.

Enfin, l'ergonomie de l'outil informatique a été améliorée pour faciliter le signalement des cas de Covid-19. En parallèle du développement par Santé publique France d'une partie de l'interface de Voozoo permettant aux ESMS l'accès à l'ensemble de leurs signalements, une prestation a été conduite pour améliorer la page d'accueil.

Fin de SurvESMS

À la suite des évolutions de la surveillance de la Covid-19 en France qui fait désormais l'objet de bulletins hebdomadaires sur les IRA, le dispositif de surveillance SurvESMS a été arrêté le 20 juin 2023, afin de l'intégrer au dispositif de signalement des épisodes groupés d'IRA dans le portail de signalement des événements sanitaires indésirables du ministère de la Santé. Cette évolution du dispositif conduit à un suivi pour la prise en charge initiale des épisodes de cas groupés de Covid-19 par les ARS, puis dans un second temps, une surveillance épidémiologique par Santé publique France.

Le signalement des IRA sur ce portail, qui concernait jusqu'à présent la grippe et le virus respiratoire syncytial (VRS) en Ehpad, propose désormais le signalement des cas groupés de Covid-19. Le signalement dès l'identification du premier cas est abandonné au profit d'un signalement lors de la survenue d'au moins trois cas d'IRA parmi les résidents dans un délai de quatre jours. Un second et dernier formulaire est renseigné en fin d'épisode, lorsqu'il s'est écoulé au moins 14 jours en l'absence de nouveau cas, à partir de la date de survenue du dernier cas dans l'établissement²¹.

Discussion-conclusion

Les limites dans l'architecture des systèmes d'information en santé, en particulier couvrant les ESMS, n'ont permis de disposer d'indicateurs nationaux de surveillance de la Covid-19 dans les ESMS qu'à

partir de la fin mars 2020²². Cependant, le dispositif de surveillance SurvESMS a été construit en urgence et a su s'adapter aux évolutions des connaissances et des pratiques relatives à la pandémie de Covid-19. Ce dispositif a permis, d'une part, le signalement des épisodes de Covid-19 par les ESMS ciblés sur l'ensemble du territoire, et d'autre part, la production des indicateurs de surveillance utiles à la gestion des épisodes, déclinés aux niveaux national, régional et départemental.

L'expérience a montré la nécessité de disposer, pour les acteurs de l'évaluation et de la gestion dans une situation de crise sanitaire, d'un système d'information réactif pour la détection et l'alerte précoce des signaux ainsi que le suivi épidémiologique s'approchant de l'exhaustivité. Ainsi, la surveillance quotidienne de la Covid-19 et la restitution journalière des données a été un enjeu majeur qui a motivé le développement spécifique du dispositif de surveillance SurvESMS. Une fois la situation de crise sanitaire finie, le signalement des épisodes de Covid-19 en ESMS a rejoint le système de signalement des IRA, dont la fréquence de signalement réduite est davantage destinée à l'alerte, mais qui n'est plus compatible avec la surveillance épidémiologique quotidienne.

L'avantage de disposer d'une plateforme informatique au sein de Santé publique France a permis d'éviter le recours à la prestation externe plus lourde à mettre en œuvre. L'agence ayant la maîtrise de la plateforme Voozanoo a pu développer et adapter rapidement les formulaires aux évolutions de la surveillance.

L'évolution du dispositif, ayant intégré dans sa deuxième version le contrôle de la qualité des données et la restitution automatisée des indicateurs, a allégé la charge de travail menée par les acteurs régionaux. Cependant, les ARS et les cellules régionales de Santé publique France sont restées très mobilisées, afin de veiller à l'exhaustivité et à la qualité des données recueillies. Des échanges se tenaient régulièrement entre les établissements, les cellules régionales et les ARS afin de s'assurer de la mise à jour des épisodes de cas de Covid-19 et de leur clôture. Ces difficultés quant à l'exhaustivité et le contrôle des données ont été rencontrées dans des systèmes de surveillance comparables²³.

La déclaration des épisodes de Covid-19 étant fondée sur le volontariat, elle ne pouvait pas être exhaustive. En période de pandémie de Covid-19, les contraintes organisationnelles qui ont pesé sur les ESMS²⁴ ont pu conduire à un défaut de signalement d'épisode de cas de Covid-19. En outre, des confusions dans le circuit de signalement ont pu être constatées. Certains épisodes étaient déclarés uniquement aux ARS sans être déclarés dans le dispositif de surveillance SurvESMS.

L'amélioration de la qualité des données a été notamment possible grâce à la capacité des établissements à corriger eux-mêmes leurs données. L'amélioration du dispositif de surveillance permettant l'accès des établissements à leurs données a été possible grâce

à l'exploitation, d'une part, de bases de données et de référentiels des établissements sanitaires, et d'autre part, d'annuaires des contacts dans les établissements.

Malgré l'évolution apportée à l'ergonomie de la solution informatique proposée, la plateforme Voozanoo 3 vieillissante manque de fonctionnalités utiles. En particulier, des difficultés de connexion liées à l'absence, l'obsolescence ou la perte des identifiants nécessaires à l'authentification des ESMS ont pu avoir un impact sur le signalement par les établissements. La charge importante du support applicatif, notamment pour le renouvellement des identifiants, a nécessité le recours à une prestation externe à Santé publique France.

Une des limites importantes du dispositif était l'absence d'interopérabilité avec d'autres systèmes d'information, et notamment le Système d'information de veille et sécurité sanitaires (SI-VSS) permettant aux ARS d'assurer la traçabilité et la gestion des signalements d'événements sanitaires dans leur région^{25,26}. L'interopérabilité entre SurvESMS et SI-VSS aurait amélioré l'articulation entre surveillance et gestion de l'épidémie et contribué notamment à réduire le délai de mise en œuvre des mesures de gestion par les ARS.

Cette capacité de développement rapide d'un dispositif de surveillance en situation d'émergence exige une préparation à la crise sur les enjeux spécifiques des systèmes d'information. La préparation requiert une réflexion sur la structure et les fonctionnalités nécessaires d'un système d'information adapté à la surveillance en période de crise sanitaire. Le développement du dispositif de surveillance SurvESMS a mis en lumière la nécessité de se doter de plateformes informatiques performantes, agiles et interopérables permettant le recueil quotidien de données de santé visant l'exhaustivité. La réflexion sur les fonctionnalités doit porter sur les modalités d'acquisition, de contrôle, de traçabilité, de traitement et de restitution des données à toutes les échelles géographiques de l'évaluation et de la gestion de crise. La préparation à la crise sanitaire nécessite également la mise à disposition de bases de données, de référentiels et d'annuaires nécessaires pour construire un tel dispositif de surveillance. Enfin, la construction d'un schéma directeur des systèmes d'information de la gestion de crise sanitaire²² ne peut être menée que dans une vision holistique des dispositifs de surveillance existants et avec la connaissance du besoin des différents acteurs concernés. ■

Remerciements

Les auteurs remercient tout d'abord l'ensemble des partenaires impliqués dans la surveillance de la Covid-19 en ESMS et en particulier les professionnels des établissements. Les auteurs remercient également toutes les personnes de Santé publique France qui ont contribué à la mise en œuvre de ce dispositif de surveillance : Kostas Danis, Scarlett Georges, Florian Franke, Karine Wyndels, Katia Hamdad, Fatima Etemadi, Didier Che, Bruno Coignard, Nicolas Methy, Bruno Hubert, Stella Laporal, Audrey Leon, Ghaya Ben Hmidene, Fanny Chereau, Camille Le Gal, Laure Fonteneau, Minh-Canh Quan, Jeanne Tamarelle,

Clémentine Calba, Amandine Cochet, Ursula Noury, Mélanie Martel, Cécile Durand, Pascaline Loury, Laurence Pascal, Myriam Blanchard, Karine Mantey, Mélanie Yvroud, Dominique Jeannel, Garance Terpent, Marion Philippe.

Liens d'intérêt

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

- [1] Applegate WB, Ouslander JG. Covid-19 presents high risk to older persons. *J Am Geriatr Soc.* 2020;68(4):681.
- [2] Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, *et al.* Clinical characteristics of 138 hospitalized patients with 2019 novel Coronavirus-infected pneumonia in Wuhan, China. *JAMA.* 2020;323(11):1061-9.
- [3] Liu K, Chen Y, Lin R, Han K. Clinical features of Covid-19 in elderly patients: A comparison with young and middle-aged patients. *J Infect.* 2020;80(6):e14-8.
- [4] Ioannou GN, Locke E, Green P, Berry K, O'Hare AM, Shh JA, *et al.* Risk factors for hospitalization, mechanical ventilation, or death among 10 131 US Veterans with SARS-CoV-2 infection. *JAMA Netw Open.* 2020;3(9):e2022310.
- [5] D'Adamo H, Yoshikawa T, Ouslander JG. Coronavirus disease 2019 in geriatrics and long-term care: The ABCDs of Covid-19. *J Am Geriatr Soc.* 2020;68(5):912-7.
- [6] Fallon A, Dukelow T, Kennelly SP, O'Neill D. Covid-19 in nursing homes. *QJM.* 2020;113(6):391-2.
- [7] White EM, Kosar CM, Feifer RA, Blackman C, Gravenstein S, Ouslander J, *et al.* Variation in SARS-CoV-2 prevalence in U.S. skilled nursing facilities. *J Am Geriatr Soc* 2020;68(10):2167-73.
- [8] Ladhani SN, Chow JY, Janarthanan R, Fok J, Crawley-Boevey E, Vusirikala A, *et al.* Increased risk of SARS-CoV-2 infection in staff working across different care homes: Enhanced Covid-19 outbreak investigations in London care homes. *J Infect.* 2020;81(4):621-4.
- [9] Fioni J, Campèse C, Spaccferri G, Rolland P, Caserio-Schönemann C, Che D. Structuration évolutive d'une surveillance multi-sources pour répondre à une infection émergente : l'expérience française face à la Covid-19. *Bull Épidémiol Hebd.* 2023;(1):2-16. http://beh.santepubliquefrance.fr/beh/2023/1/2023_1_1.html
- [10] Ministère des solidarités et de la santé. Direction générale de la santé. Centre de crise sanitaire. MINSANTE/CORRUSS n° 2020_46 : Signalement des cas Covid-19 dans les établissements sociaux et médico-sociaux (EMS).
- [11] Santé publique France. Signalement de cas de Covid-19 dans les établissements sociaux et médico-sociaux – Guide pour les établissements. Saint-Maurice: Santé publique France; 2021. 40 p. https://www.santepubliquefrance.fr/content/download/327410/file/Guide_Signalement_ESMS_COVID-19_ESMS_20210305.pdf
- [12] Société française d'hygiène hospitalière. Actualisation des précautions standard – Établissements de santé, Établissements médicaux sociaux, Soins de ville. *Hygiènes.* 2017;25 (Hors-série):1-68. https://www.sf2h.net/k-stock/data/uploads/2017/06/HY_XXV_PS_versionSF2H.pdf
- [13] Société française d'hygiène hospitalière. Prévention de la transmission croisée par voie respiratoire : Air ou Gouttelettes. Recommandations pour la pratique clinique (RPC). *Hygiènes.* 2013;20(1):1-60. <https://www.sf2h.net/publications/prevention-de-la-transmission-croisee-par-voie-respiratoire-air-ou-gouttelettes.html>
- [14] Société française d'hygiène hospitalière. Prévention de la transmission croisée : précautions complémentaires contact. Recommandations nationales. *Hygiènes.* 2009;17(2):81-138.

<https://www.sf2h.net/publications/prevention-de-la-transmission-croisee-precautions-complementaires-contact.html>

- [15] Ministère des Solidarités et de la Santé. Instruction N° DGS/VSS1/DGCS/SPA/2019/211 du 30 septembre 2019 relative aux conduites à tenir devant des infections respiratoires aiguës ou des gastro-entérites aiguës dans les collectivités de personnes âgées. 2019. 9p. <https://www.preventioninfection.fr/document/instruction-n-dgs-vss1-dgcs-spa-2019-211-du-30-septembre-2019-relative-aux-conduites-a-tenir-devant-des-infections-respiratoires-aigues-ou-des-gastro-enterites-aigues-dans-les-collectivites-de-p/>
- [16] Santé publique France. Covid-19 Point épidémiologique hebdomadaire du 25 juin 2020. Surveillance dans les établissements sociaux et médico-sociaux (ESMS). Saint-Maurice: Santé publique France; 2020. pp. 10-11. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-25-juin-2020>
- [17] Belmin J, Georges S, Franke F, Daniau C, Cochet A, Durand C, *et al.* Coronavirus disease 2019 in French residential care facilities: A nationwide study. *J Am Med Dir Assoc.* 2021;22(6):1142-5.
- [18] ECDC Public Health Emergency Team, Danis K, Fonteneau L, Georges S, Daniau C, Bernard-Stoecklin S, *et al.* High impact of Covid-19 in long-term care facilities, suggestion for monitoring in the EU/EEA, May 2020. *Euro Surveill.* 2020;25(22):2000956. <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.22.2000956>
- [19] Sanchez Ruiz MA, Franke F, Giron S, Boulogne O, Aventini C, Marchand E, *et al.* Le fardeau de la Covid-19 dans les Ehpad de Provence-Alpes-Côte d'Azur, 2020-2022. *Bull Épidémiol Hebd.* 2023;(18):370-9. http://beh.santepubliquefrance.fr/beh/2023/18/2023_18_3.html
- [20] Miron de l'Espinay A, Ricroch L. En 2020, trois Ehpad sur quatre ont eu au moins un résident infecté par la Covid-19. *Études et Résultats.* 2021;(1196):1-8. <https://drees.solidarites-sante.gouv.fr/publications/etudes-et-resultats/en-2020-trois-ehpad-sur-quatre-ont-eu-au-moins-un-resident-infecte>
- [21] Santé publique France. Signalement des épisodes de cas groupés d'infection respiratoire aiguë (IRA) dans les établissements médico-sociaux (EMS). Guide pour les établissements. Saint-Maurice: Santé publique France; 2023. 25 p. <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/documents/guide/signalement-des-episodes-de-cas-groupes-d-infection-respiratoire-aigue-ira-dans-les-etablissements-medico-sociaux-ems-.guide-pour-les-etablisse>
- [22] Pittet D, Boone L, Moulin AM, Briet R, Parneix P. Mission indépendante nationale sur l'évaluation de la gestion de la crise Covid-19 et sur l'anticipation des risques pandémiques – Rapport final. 2021. 179 p. <https://www.vie-publique.fr/rapport/279851-gestion-crise-covid-et-anticipation-des-risques-pandemiques-rapport-final>
- [23] Gossin M, Walther D, Blanco JM, Masserey É, Meylan L, Pittet V, *et al.* SICOVID : un système cantonal d'information COVID pour la décision en santé publique. *Rev Med Suisse.* 2020;16:2177-82. <https://www.revmed.ch/revue-medicale-suisse/2020/revue-medicale-suisse-714/sicovid-un-systeme-cantonal-d-information-covid-pour-la-decision-en-sante-publique>
- [24] Gonthier R, Adolphe M, Michel JP, Bringer J, Dubois B, Lecomte D, *et al.* Rapport 22-02. Après la crise Covid, quelles solutions pour l'Ehpad de demain ? *Bull Acad Natl Med.* 2022;206:457-65. <https://www.academie-medecine.fr/apres-la-crise-covid-quelles-solutions-pour-lehpad-de-demain/>

[25] Ministère des Affaires sociales et de la Santé. Instruction DGS/DUS/CORRUSS n° 2013-274 du 27 juin 2013 relative à l'organisation territoriale de la gestion des situations sanitaires exceptionnelles. 2013. 17 p. https://sante.gouv.fr/fichiers/bo/2013/13-08/ste_20130008_0000_0176.pdf

[26] Desenclos JC. Les acteurs de l'alerte. ADSP. 2019;106:22-5.

Citer cet article

Daniau C, Naud J, Tessier S, Hassan E. SurvESMS : un dispositif de surveillance de la Covid-19 en établissements sociaux et médico-sociaux. Bull Épidémiol Hebd. 2024;(20-21):481-8. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_6.html

> FOCUS // Focus

LES SYSTÈMES D'INFORMATION À L'ÉPREUVE DE LA COVID-19 : ENSEIGNEMENTS, NOUVEAUX ENJEUX ET PERSPECTIVES POUR SE PRÉPARER AUX PROCHAINES CRISES

// INFORMATION SYSTEMS PUT TO THE TEST BY COVID-19: LESSONS LEARNED, NEW CHALLENGES AND OPPORTUNITIES TO PREPARE FOR FUTURE CRISES

Céline Caserio-Schönemann¹ (celine.caserio-schonemann@santepubliquefrance.fr), Adel Arfaoui¹, Bruno Coignard¹, Yann Le Strat¹, Patrick Rolland¹, Guillaume Spaccaverri²

¹ Santé publique France, Saint-Maurice

² Santé publique France – Auvergne-Rhône-Alpes, Lyon

Soumis le 02.08.2024 // Date of submission: 08.02.2024

Mots-clés : Systèmes d'information, Covid-19, Surveillance épidémiologique, Crise sanitaire
// **Keywords:** Information systems, COVID-19, Epidemiological surveillance, Health crisis

L'épidémie de Covid-19 a montré l'importance de s'appuyer sur un système multisource territorialisé de surveillance épidémiologique, permettant de suivre de façon réactive et continue la dynamique d'un phénomène émergent, pour mieux contrôler sa diffusion dans la population. Le dispositif de surveillance multisource de la Covid-19, déployé par Santé publique France, s'est ainsi adossé à tous les secteurs du champ de la santé, en interface avec de multiples systèmes d'information (SI) et partenaires : la médecine de ville (SOS Médecins, réseau Sentinelles), les laboratoires de biologie médicale (SI-DEP⁽¹⁾ et Emergen), les pharmacies (SI-DEP), les structures de vaccination (VAC-SI⁽²⁾), l'hôpital dans toutes ses composantes : les urgences (Oscour[®]⁽³⁾), les services conventionnels et de soins critiques (SI-VIC⁽⁴⁾ et surveillance des cas graves en réanimation), y compris en pédiatrie (surveillance en néonatalogie, Pandor⁽⁵⁾, Picure⁽⁶⁾), les établissements sociaux et médico-sociaux dont les Ehpad⁽⁷⁾ (SurvESMS⁽⁸⁾), le suivi des vaccinations (VAC-SI), la surveillance des

clusters (Monic⁽⁹⁾), le suivi des contacts (Cnam⁽¹⁰⁾), la surveillance de la santé mentale¹ (Oscour[®], SOS Médecins, enquêtes Coviprev...) et la mortalité (Insee⁽¹¹⁾, CépiDc-Inserm⁽¹²⁾, SI-VIC, ESMS)².

En ces temps exceptionnels de pandémie, il a été demandé à Santé publique France de généraliser la fréquence quotidienne de production et de restitution de tous les indicateurs ; cette fréquence proche du réel est apparue comme répondant aux besoins d'information des autorités et de communication vers la population générale quant à la diffusion de l'épidémie sur le territoire. Hors de ce contexte inédit, la temporalité de restitution des indicateurs de surveillance doit être définie au cas par cas lors de la mise en œuvre de chaque nouvelle surveillance, en fonction des objectifs poursuivis, des caractéristiques épidémiologiques de la maladie ou du phénomène surveillé, et pour l'adaptation ou la mise en place efficiente de mesures de gestion au niveau populationnel. Elle devra dans tous les cas s'appuyer sur des SI robustes, intégrés aux dispositifs métiers des partenaires, de conception agile et optimisée, et suffisamment dimensionnés.

Le travail intense qui a été nécessaire pour mettre en place, adapter et/ou articuler les SI sur lesquels s'est appuyé tout le dispositif de surveillance de la Covid-19

⁽¹⁾ SI-DEP : Système d'information de dépistage populationnel.

⁽²⁾ VAC-SI : Système de suivi des vaccinations.

⁽³⁾ Oscour[®] : Organisation de la surveillance coordonnée des urgences.

⁽⁴⁾ SI-VIC : Système d'information pour le suivi des victimes.

⁽⁵⁾ Pandor : réseau de surveillance des hospitalisations des enfants en pédiatrie générale.

⁽⁶⁾ Picure : réseau de surveillance des hospitalisations des enfants en réanimation pédiatrique.

⁽⁷⁾ Ehpad : Établissements d'hébergement pour personnes âgées dépendantes.

⁽⁸⁾ SurvESMS : dispositif de surveillance de la Covid-19 dans les établissements sociaux et médico-sociaux.

⁽⁹⁾ Monic : Monitoring des clusters.

⁽¹⁰⁾ Cnam : Caisse nationale de l'assurance maladie.

⁽¹¹⁾ Insee : Institut national de la statistique et des études économiques.

⁽¹²⁾ CépiDc-Inserm : Centre d'épidémiologie sur les causes médicales de décès de l'Institut national de la santé et de la recherche médicale.

a été l'occasion de faire émerger des constats sur les avantages, limites et marges de progression pour la construction des systèmes d'information pour la surveillance de demain.

Pour être pleinement opérationnels en cas de crise, les SI en santé doivent idéalement :

- être déjà fonctionnels hors crise pour les cas d'urgence, afin d'éviter les écueils d'une mise en place en urgence, non anticipée ;
- être déployés sur l'ensemble du territoire pour répondre aux besoins de tous ;
- associer les producteurs de données de manière précoce et continue, pour faciliter leur adhésion aux dispositifs mis en place ;
- permettre, autant que possible, de collecter des données déjà renseignées par les professionnels de santé pour leurs propres besoins métiers, en évitant des saisies manuelles redondantes et chronophages qui, outre la surcharge de travail, rendent la stratégie poursuivie difficilement lisible ;
- être interopérables avec des SI déjà existants ;
- être suffisamment flexibles et évolutifs, pour s'adapter à l'évolution du phénomène surveillé et tenir compte de l'acquisition progressive des connaissances ;
- être complémentaires avec les dispositifs existants pour permettre de couvrir les différents risques, qu'ils soient d'origine infectieuse, environnementale, professionnelle, technologique, industrielle ou malveillante ;
- couvrir également les différents secteurs de prise en charge des patients pour prendre en compte l'évolution possible des modalités du recours aux soins, que ce soit du fait de l'évolution de la maladie (gravité, nouvelle présentation clinique...) ou de celle de l'organisation de la prise en charge ;
- permettre, via un identifiant unique, de croiser les bases de données entre elles tout en garantissant la protection des données sensibles et la non ré-identification des personnes ;
- s'appuyer sur une co-construction par les différents acteurs et organismes producteurs et utilisateurs, pour mutualiser les ressources en vue de répondre à un panel d'objectifs complémentaires (donc éviter des systèmes parallèles et redondants) ;
- disposer de ressources suffisantes, aussi bien techniques (infrastructure adaptée, hautes capacités de calcul et de stockage) qu'humaines (redondance indispensable des fonctions pour assurer la continuité d'activité, compétences techniques appropriées). L'un des enseignements majeurs de la crise a été de mettre en évidence le besoin de nouvelles compétences en *data engineering* et *data science* ;

- être innovants, tant en termes de technologies utilisées que d'organisation pour les mettre en œuvre et en assurer la continuité en toute circonstance ;
- préparer le cadre et les conditions de la pérennisation des dispositifs ayant fait la preuve de leur utilité et de leur efficacité après la crise, une fois l'état d'urgence terminé.

Certains dispositifs comme le système de surveillance syndromique SurSaUD® (Surveillance sanitaire des urgences et des décès) disposaient déjà de certaines de ces caractéristiques avant la crise de la Covid-19^{3,4}. Aujourd'hui, d'autres SI en cours de développement depuis la fin de la crise proposent de tester la production réactive d'indicateurs à partir des entrepôts de données de santé des établissements hospitaliers (EDSH), permettant de conserver la donnée source dans son environnement de production sans avoir à gérer les contraintes liées à la mise en place de flux sécurisés, comme le projet Orchidée (Organisation d'un réseau de centres hospitaliers impliqués dans la surveillance épidémiologique et la réponse aux émergences).

La mise à disposition en open data des indicateurs produits est nécessaire pour améliorer la transparence, le partage et la ré-utilisation, dans le cadre d'une ouverture plus large vers la société. Les outils mis en place par l'agence, avant puis pendant la pandémie (en particulier l'observatoire cartographique Géodes⁽¹³⁾ avec plus de 800 indicateurs mis à disposition), sont actuellement en cours d'évolution pour mieux répondre aux besoins des différents publics. Cette mise à disposition doit se faire à des niveaux géographiques adaptés pour la prise de décision par les autorités sanitaires régionales. Publiés à l'échelle de l'intercommunalité (EPCI), ils peuvent permettre également à chaque citoyen qui le souhaite, de mieux accéder à une information locale qui le concerne directement. L'accompagnement de cette information pour la rendre compréhensible et utile est indispensable, ce qui implique que les efforts doivent autant se porter sur le type et le nombre d'indicateurs à publier que sur les éléments permettant aux citoyens de développer leur littératie, afin de favoriser leur adhésion aux mesures prises par les autorités et leur confiance dans l'action publique.

Dans le domaine des SI, la France doit s'organiser aujourd'hui pour s'attacher à clarifier au maximum les responsabilités respectives (notamment pour le *contact-tracing*) et les champs encore insuffisamment couverts (médecine de ville, inégalités sanitaires et sociales, personnes âgées dépendantes...). Certains chantiers ont pu évoluer grâce à l'opportunité de la Coupe du monde de rugby 2023 et des Jeux olympiques et paralympiques (JOP) de Paris 2024, comme la remontée des données des Samu (Service d'aide médicale urgente) à des fins de veille sanitaire et de surveillance épidémiologique, avec le soutien de l'Agence du numérique en santé. Ce nouveau

⁽¹³⁾ <https://geodes.santepubliquefrance.fr/>

dispositif devra être évalué après les JOP pour envisager son éventuelle pérennisation, comme héritage des Jeux. D'autres chantiers structurants pour la surveillance nécessiteraient d'être poursuivis, renforcés, voire développés pour certains, en s'appuyant sur un engagement volontariste et une collaboration active des structures et organismes impliqués :

- l'accélération du déploiement de la certification électronique des décès ;
- l'élargissement de la remontée des tests biologiques de la Covid-19 à d'autres agents pathogènes, y compris non infectieux, via le système d'information Laboé-SI qui devrait prendre à terme le relai de SI-DEP ;
- la création d'un SI national dans les établissements hospitaliers, incluant les services de réanimation. Le projet Orchidée devrait permettre, sur la base d'un prototype construit pour la surveillance des infections respiratoires aiguës basses (IRA) à l'automne 2024, de proposer un modèle à déployer pour d'autres pathologies, y compris non infectieuses ;
- la création d'un SI national dans les établissements sociaux et médico-sociaux (ESMS) ;
- la pérennisation d'un SI commun permettant le partage des données de génomique entre acteurs de la surveillance et de la recherche et son élargissement progressif à d'autres pathogènes infectieux susceptibles d'être à l'origine de prochaines émergences ou pandémies (Emergen 2.0) ;
- l'extension du dispositif de surveillance microbiologique des eaux usées (Sum'Eau) à d'autres pathogènes que le SARS-CoV-2 ;
- la mise à disposition d'un SI réactif pour l'enregistrement sécurisé de tous les actes de vaccination de la population française ;
- dépassant le strict champ des SI, la réflexion sur l'intégration d'indicateurs issus d'autres secteurs (agriculture, alimentation, biodiversité, comportements...) qui pourraient alimenter la démarche de refonte de l'open data menée actuellement. Cette restitution multisource et multichamp permettrait de contribuer à se préparer aux nouvelles menaces, au croisement des problématiques infectieuses et environnementales (comme les arboviroses par exemple),

en santé humaine comme animale au sein d'écosystèmes, dans le cadre d'une démarche *One Health*.

Ces axes de développement SI sont parmi les plus essentiels pour que Santé publique France puisse renforcer son dispositif de surveillance multisource et continuer à produire des indicateurs robustes et de qualité dans tous les secteurs et sur l'ensemble des territoires. Il apparaît crucial de considérer l'éthique en santé publique comme une préoccupation institutionnelle intégrée à la construction des SI, sur lesquels s'adosseront les systèmes de surveillance du futur, dans une démarche d'ouverture, de dialogue et de co-construction avec la société. L'effort est conséquent, mais il doit être mené dès à présent pour que le pays soit prêt à répondre efficacement aux prochaines crises. ■

Liens d'intérêts

Les auteurs déclarent ne pas avoir de liens d'intérêt au regard du contenu de l'article.

Références

- [1] Du Roscoät E, Forgeot C, Léon C, Doncarli A, Pirard P, Tebeka S, *et al*. La santé mentale des Français pendant l'épidémie de Covid-19 : principaux résultats de la surveillance et des études conduites par Santé publique France entre mars 2020 et janvier 2022. *Bull Épidémiol Hebd*. 2023;(26):570-89. http://beh.santepubliquefrance.fr/beh/2023/26/2023_26_2.html
- [2] Figoni J, Campèse C, Spaccaferri G, Rolland P, Caserio-Schönemann C, Che D. Structuration évolutive d'une surveillance multi-sources pour répondre à une infection émergente : l'expérience française face à la Covid-19. *Bull Épidémiol Hebd*. 2023;(1):2-16. http://beh.santepubliquefrance.fr/beh/2023/1/2023_1_1.html
- [3] Thiam MM, Pontais I, Forgeot C, Pedrono G, SurSaUD® Regional Focal Point, SOS Médecins, *et al*. Syndromic surveillance: A key component of population health monitoring during the first wave of the COVID-19 outbreak in France, February-June 2020. *PLoS One*. 2022;17(2):e0260150.
- [4] Olié V, Carcaillon-Bentata L, Thiam MM, Haeghebaert S, Caserio-Schönemann C. Emergency department admissions for myocardial infarction and stroke in France during the first wave of the COVID-19 pandemic: National temporal trends and regional disparities. *Arch Cardiovasc Dis*. 2021;114(5):371-80.

Citer cet article

Caserio-Schönemann C, Arfaoui A, Coignard B, Le Strat Y, Rolland P, Spaccaferri G. Les systèmes d'information à l'épreuve de la Covid-19 : enseignements, nouveaux enjeux et perspectives pour se préparer aux prochaines crises. *Bull Épidémiol Hebd*. 2024;(20-21):488-90. http://beh.santepubliquefrance.fr/beh/2024/20-21/2024_20-21_7.html